JAN 5 1982

1051- H-11

Y3. N88: 25/2372

NUREG/CR-2372

Properties of Parameter Estimation Techniques for a Beta-Binomial Failure Model

ORI

COMPLETED

Prepared by J. K. Shultis, W. Buranapan, N. D. Eckhoff

Department of Nuclear Engineering
Kansas State University

Prepared for
U.S. Nuclear Regulatory
Commission

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

Available from

GPO Sales Program
Division of Technical Information and Document Control
U. S. Nuclear Regulatory Commission
Washington, D. C. 20555

Printed copy price: $6.00

and

National Technical Information Service
Springfield, Virginia 22161

NUREG/CR-2372
KSU-2075b

Properties of Parameter Estimation Techniques for a Beta-Binomial Failure Model

Manuscript Completed: August 1981
Date Published: December 1981

Prepared by
J. K. Shultis, W. Buranapan, N. D. Eckhoff

Department of Nuclear Engineering
Kansas State University
Manhattan, KS 66506

Prepared for
Division of Risk Analysis
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
NRC FIN A6290

ABSTRACT

To use the compound beta-binomial probability model in the analysis of component failure data of the failure-on-demand type, specific values are first needed for the parameters of the beta function which serves as the prior distribution of component failure probabilities. Five methods for estimating the beta parameters from observed failure data are examined in this work: (i) matching moments of the prior distribution to data moments, (ii) maximum likelihood method based on the prior distribution, (iii) weighted marginal matching moments, (iv) unweighted marginal matching moments, and (v) maximum likelihood method based on the marginal distribution. The distribution of the beta estimators for each method was obtained empirically by using various sized samples of simulated failure data that were randomly generated from a known beta-binomial distribution skewed towards low failure probabilities.

From these empirical distributions, many properties of the estimators were examined. It was found that the prior matching moments method, which is computationally the simplest and which almost always yield parameter estimates, gives the smallest bias and mean squared error in the parameter estimates for small sample sizes (≤ 10). This method also yields estimated prior distributions which are more conservative from a safety viewpoint than those obtained by the other estimation methods. Moreover, the other estimation methods occasionally fail either to give any parameter estimates or to produce realistic parameter estimates from failure data of small sample size. For large sample sizes (≥ 20), the marginal maximum likelihood estimation method becomes superior while the prior maximum likelihood method amost always fails to give any parameter estimators and hence is judged unsuitable for the analysis of failure data from components with low failure probabilities.

BLANK PAGE

TABLE OF CONTENTS

Page

LIST OF FIGURES vii

LIST OF TABLES ix

1.0 INTRODUCTION. 1
  1.1 Statistical Failure Models 2
  1.2 Models with a Beta Prior Distribution. 7
  1.3 Construction of Prior Distribution 11
  1.4 Scope of Study 13

2.0 PRIOR PARAMETER ESTIMATION TECHNIQUES 16
  2.1 Method of Matching Moments of the Prior Distribution to Data Moments 17
  2.2 Maximum Likelihood Method Based on the Prior Distribution 21
  2.3 Method of Weighted Marginal Matching Moments 26
  2.4 Method of Unweighted Marginal Matching Moments 29
  2.5 Maximum Likelihood Method Based on the Marginal Distribution 30
  2.6 Lower Bounds on the Variances of Prior Parameter Estimates. 33

3.0 SIMULATION STUDY OF PRIOR ESTIMATION TECHNIQUES 38
  3.1 Desirable Properties of an Estimator 38
  3.2 Generation of Simulated Failure Data 40

4.0 RESULTS AND DISCUSSION. 46
  4.1 Distribution of Prior Parameter Estimators 47
    4.1.1 Mean and Variance 47
    4.1.2 Bias of Prior Parameter Estimates 63
    4.1.3 Mean Squared Error of Estimators. 65
    4.1.4 Median of Estimators. 67
    4.1.5 Comparison to Results From a Symmetric Beta Prior. 70
  4.2 Distribution of Estimators for the Mean and Variance of the Prior Distribution. 77
  4.3 Distribution of 95th Percentile Estimators 84
  4.4 Fraction of the Estimated Prior Distribution Above the True 95-th Percentile. 98
  4.5 Comparison of Maximum Likelihood Variance Bounds to Measured Variances. 105

5.0 CONCLUSIONS 112

6.0 REFERENCES. 115

Table of Contents - Continued

Page

8.0 APPENDICES. 116

APPENDIX A: Derivation of Expectation of Statistic S of Eq. (2.20). 117

APPENDIX B: Evaluation of Polygamma Functions 120

APPENDIX C: Evaluation of the Incomplete Beta Functions 122

APPENDIX D: Computer Program Listings 124

LIST OF FIGURES

Figure		Page
4.1	Distribution of beta parameter estimators for samples of size N=5.	53
4.2	Distribution of beta parameter estimators for samples of size N=10	55
4.3	Distribution of beta parameter estimators for samples of size N=20	57
4.4	Distribution of beta parameter estimators for samples of size N=50	59
4.5	Variation of the bias of the beta parameter estimators with sample size for the different estimation techniques (a=1.2, b=23)	66
4.6	Variation of the mean squared error of the beta parameter estimators with sample size for the different estimation techniques (a=1.2, b=23).	69
4.7	Variation of the median of the beta parameter estimators with sample size for the different estimation techniques (a=1.2, b=23).	72
4.8	Variation of the bias of the beta parameter estimators with sample size for the symmetric beta distribution (a=b=5)	74
4.9	Variation of the mean squared error of the beta parameter estimators with sample size for the symmetric beta distribution (a=b=5).	75
4.10	Distribution of the means of the estimated beta prior distributions from samples of size N=5	79
4.11	Distribution of the means of the estimated beta prior distributions from samples of size N=10.	80
4.12	Distribution of the means of the estimated beta prior distributions from samples of size N=20.	81
4.13	Distribution of the means of the estimated beta prior distributions from samples of size N=50.	82
4.14	Distribution of the variances of the estimated beta prior distributions from samples of size N=5.	85

List of Figures - Continued

Figure Page

4.15 Distribution of the variances of the estimated beta prior distributions from samples of size N=10. . . 86

4.16 Distribution of the variances of the estimated beta prior distributions from samples of size N=20. . . 87

4.17 Distribution of the variances of the estimated beta prior distributions from samples of size N=50. . . 88

4.18 Distribution of 95-th percentiles of the estimated beta prior distributions for samples of size N=5. . . . 91

4.19 Distribution of 95-th percentiles of the estimated beta prior distributions for samples of size N=10 . . . 92

4.20 Distribution of 95-th percentiles of the estimated beta prior distributions for samples of size N=20 . . . 93

4.21 Distribution of 95-th percentiles of the estimated beta prior distributions for samples of size N=50 . . . 94

4.22 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data (a=1.2,b=23) for sample size N=5. 101

4.23 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data (a=1.2, b=23) for sample size N=10. 102

4.24 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data (a=1.2, b=23) for sample size N=20. 103

4.25 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data (a=1.2, b=23) for sample size N=50. 104

LIST OF TABLES

Table Page

3.1 Algorithm for optimal calculation of number of failures, k, by the inverse transformation technique 45

4.1 Number of successful solutions and failures for prior parameter estimates from the simulation failure data for the five estimation techniques (a=1., b=23) 48

4.2 Simulated failure data $\binom{k}{n}$ from a beta-binomial (a=1.2, b=23) for which the marginal-based estimation methods yielded no solution 49

4.3 Mean and number of solutions of parameter estimates for different sample sizes and estimation techniques (a=1.2, b=23) 61

4.4 Variances and covariance of parameter estimators for different sample sizes and estimation techniques (a=1.2, b=23) 62

4.5 The bias or deviation of mean of estimators from true parameters (a=1.2, b=23) 64

4.6 Mean squared error about the true beta parameters (a=1.2, b=23) for the simulated failure data 68

4.7 Median values for the estimates $\hat{a}$ and $\hat{b}$ (a=1.2, b=23) for different sample size and estimation techniques 71

4.8 The bias and mean squared error of the estimators of the parameters for a symmetric beta prior distribution (a=b=5) as calculated by different estimation techniques from simulated failure data of various sample sizes 76

4.9 Mean and variance of the estimators for the mean of the beta prior (a=1.2, b=23) for different sample sizes 83

4.10 Mean and variance of the estimators for the variance of the beta prior (a=1.2, b=23) for different sample sizes 89

4.11 Median, mean and variance of the distributions of the 95-th percentile estimators 95

List of Tables - Continued

Table Page

4.12 Number and percent of simulated failure data samples which yielded estimated 95-th percentiles greater than or less than the true value of 0.13586. 97

4.13 Smallest 95-th percentile estimators observed for simulated failure data samples of size N 99

4.14 Median, mean and variance of the distribution for the Prob{$p \geq p_{95}^{true}$} for different sample sizes (a=1.2, b=23). 106

4.15 Variance bounds, and the covariance bounds for parameter estimators, as calculated by the marginal maximum likelihood method for selected simulated failure data samples 108

4.16 Selected simulated failure data samples used to estimate variance and covariance bounds in Table 4.15 109

4.17 Ratio of measured variances and covariances of the parameter estimators to the marginal maximum likelihood bounds for the "good" data samples. 110

1.0 INTRODUCTION

The principal purpose in the development of failure models is to make predictions about the future behavior of the system under scrutiny. The better the agreement between a model and the past failure data for the system, the better one feels about his description of the stochastic nature of the failure mechanism, particularly when the development of the failure model is based on reasonable physical assumptions. Many different statistical failure models can be proposed which are based on different assumptions for the failure mechanism such as, type of failures, relation of one failure to another, the dependence of one part of the system on another, the different time evolution of the system parameters, etc.

Of considerable importance in the safety analysis of nuclear power plants are methods to estimate the probability of failure-on-demand, p, of a plant component that normally is inactive and that may fail when activated or stressed (e.g., an emergency diesel generator). The probability of failure-on-demand for such a component is defined as the probability that it will fail to perform its designed task when needed. To evaluate this failure probability for a particular component, some kind of experimental data from that component is generally needed. For example, one may test a standby emergency diesel generator by trying to start it n times during which the generator fails to start k times. The number of attempts and the observed number of failures to start give a type of attribute test data which can then be used to estimate the failure probability of the diesel generator.

From this type of attribute failure data, the failure probability of a component can be computed in a variety of ways depending on the type of statistical failure rate model one wishes to use to describe the actual physical system. Several models are described in the next two sections.

1.1 Statistical Failure Models

(a) Homogeneous Model

In the standard classical description of a component, the failure probability is usually regarded as an unknown constant and the outcome of each attempt to activate the component (e.g., start the diesel generator) is assumed to be independent of any other attempt. Thus the probability of obtaining k failures in n testings is given by the binomial distribution

$$f(k|p;n) = \binom{k}{n} p^{k}(1-p)^{n-k} , \; k=0,1,2,\ldots,n \qquad (1.1)$$

Often several similar components are treated as a single class and are assumed to have the same failure probability p. Thus repeated demands on the same component is equivalent to the same number of separate demands on different components in the class if all components are further assumed to be independent. These assumptions then yield Eq. (1.1) as a failure rate model which is often termed the "homogeneous model." By applying the maximum likelihood method to the binomial distribution [1], an estimator of failure probability $\hat{p}$, is obtained as the ratio of the total number of failures over the number of tries, i.e., k/n. Only data derived from test or field experience on the particular component of interest is used to calculate this classical estimator [2].

For a component, which is intentionally designed to have a high probability of performing its assigned task when needed, very often no failure is observed, particularly for a small number of demands. From such data, standard classical statistical methods would estimate the expected failure probability of that component to be zero which, for many applications, is unrealistic. To obtain more data and hence a better estimate of the failure probability, the failure data from similar components are often lumped together and a homogeneous model assumed. For this homogeneous model, the maximum likelihood estimator of the failure probability is simply the ratio of the sum of all failures to the sum of all demands for all components in the class. However, some writers have mentioned that in their experience such grouped failure data may exhibit a greater variation than would be expected from the homogeneous binomial model [3]. Therefore, alternatives to the simple homogeneous classical method are needed to give more reasonable results for components with inherent low failure probability characteristics.

(b) Compound Model

A more sophisticated failure model, to which the bulk of this work is devoted, is one in which the failure probability p among components lumped into a single class, is itself regarded as a random variable with a distribution g(p); i.e., the parameter p is assumed to be constant for each component in the class but will have different values from component to component. The distribution g(p) is often termed the "prior" distribution since this probability density function for p is usually obtained by previous knowledge of the components in the class under consideration.

If g(p) is known, then the joint probability density of obtaining a component from the class with a failure probability in unit p about p and which will experience k failures in n tries is [4]

$$\mathrm{Prob}\{p,k;n\} = f(k|p;n)\, g(p) \;, \quad \begin{matrix} 0<p<1 \\ k=0,1,2,\ldots,n \end{matrix} \tag{1.2}$$

where the "conditional" distribution f is given by the binomial distribution of Eq. (1.1). Upon integration over all possible values of p, one then obtains the probability that a component randomly selected from the class will experience k failures in n demands, i.e.,

$$h(k;n,g) = \int_0^1 f(k|p;n)\, g(p)\, dp, \quad k=0,1,\ldots,n \;. \tag{1.3}$$

This discrete compound distribution is termed the "marginal" distribution and is the basis for the compound model used to describe a class of similar components but each of which may have different failure probabilities. Should each component in the class have the same failure probability p_o, then the prior distribution becomes a delta function, i.e., $g(p) = \delta(p-p_o)$, and the compound model of Eq. (1.3) reduces to the homogeneous model of Eq. (1.1).

With the compound model of Eq. (1.3), one may obtain a closely related distribution. The distribution of p for a component randomly selected from the class described by g(p) and for which k failures in n demands are observed is

$$\xi(p|k;n) = g(p)\, f(k|p;n)/h(k;n,g), \; 0<p<1 \;. \tag{1.4}$$

This "posterior" distribution on p combines both the prior information g(p) and the failure data (k,n). From this posterior distribution one could then estimate the probable range (probability interval) for the failure parameter p of a component which is known to belong to a given prior class and which has experienced k failures in n demands.

(c) Use of the Compound Model in Bayesian Analysis

The insertion of extraneous information is the cornerstone of the Bayesian method [5] which over the past few years has been increasingly used in the analysis of components with low failure probabilities. In a Bayesian analysis, the failure probability p is also treated as a random variable characterized by a prior probability distribution. The prior distribution may be used to describe both the physical variation of p among components as well as the uncertainty of the analyst about what the actual values of p are. Before any actual experimental data are observed for a component, the prior distribution provides some information about the expected failure probability of the component. The prior distribution is thus an *a priori* probability model for the failure probability. This distribution represents the totality of the analyst's prior knowledge and assumptions about a component's failure probability and it should reflect the analyst's beliefs concerning the likely values of the failure probability before the observed data are obtained. Some methods which are used to construct the prior distribution will be discussed in Section 1.3.

In the standard application of Bayesian methods, the prior distribution represents a subjective judgement or bias towards a component. As such, the prior distribution can be thought of as the analyst's uncertainty or prejudice for a particular component. It is this combining of subjective opinion (or bias) with deterministic probability models (e.g., the binomial distribution) that has done much to initiate the controversy between Bayesian and classical techniques.

The classical analyst, on the other hand, wishes to avoid all subjective aspects in his probabilistic models and to use only observed failure data. Nevertheless, a classical analyst can also use prior information to improve his predictions about a particular component with a low failure probability. Rather than treat each component in isolation, he may assume (subjective?) that the component of concern belongs to a class of similar components whose constant failure probabilities are distributed according to some (prior) distribution. Each component in this class is still assumed to have a constant failure probability. One may then combine the prior information (the distribution of failure probabilities in the class to which the component of concern belongs) with the observed data for the component in question. Such was the basis for the compound model developed in the previous section.

In both the classical compound model analysis and the Bayesian analysis information about the distribution of the failure probability must be obtained. In a Bayesian analysis this information is often obtained from expert opinion, while in a classical compound model this information is given either *a priori* or must be established from data observed for the components in the class. In both analyses the subsequent development of the failure probability model for a particular component is identical to that described in the previous section. In summary, the difference between the Bayesian approach and the compound model of classical analysis is the interpretation attached to the prior distribution and the methods used to estimate it.

1.2 Models with a Beta Prior Distribution

To obtain the marginal or posterior distribution for the compound failure model (Eqs. (1.3) or (1.4)), the explicit form of the prior distribution must be known. The parameter of interest in this work is a failure probability whose possible values range from 0 to 1; therefore, in principle, any probability distribution over the interval [0,1] could serve as a prior model. Since the prior distribution in a Bayesian framework represents the analyst's beliefs about the possible values of a failure probability and different people may have different prior information or different opinions, the prior distribution model should be able to accommodate a wide variety of shapes, dispersion, mean values, etc., so as to be able to represent a wide variety of states of prior information. Similarly in a classical compound model analysis, the prior distribution must be able to assume many different shapes so as to be applicable to many different types of components.

A major difficulty in the application of the compound model is to carry out the integration in Eq. (1.3) to obtain explicitly the marginal distribution. If f and g are not fairly simple mathematical functions.it may be quite difficult to evaluate the integral of Eq. (1.3) even numerically. This integration ideally should be carried out analytically; otherwise, numerical evaluation with its inherent errors would be necessary. Because of this potential difficulty, "conjugate" prior distributions, which are families of prior distributions that ease the computational burden when used with particular conditional distributions are widely used in a compound model analysis. In addition, use of a conjugate prior distribution results in a posterior distribution which is also a member of the same conjugate family. This property is the basis for the term conjugate.

For the failure-on-demand problem, the conditional distribution is specified to be the binomial distribution. The conjugate prior for the binomial distribution is the beta distribution [6,7],

$$g(p;a,b) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}, \quad a,b \geq 0, \quad 0 \leq p \leq 1, \tag{1.5}$$

where

$$B(a,b) \equiv \int_0^1 x^{a-1}\,(1-x)^{b-1}\,dx \equiv \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \tag{1.6}$$

with the gamma function defined as

$$\Gamma(y) \equiv \int_0^\infty u^{y-1}\,e^{-u}\,du, \qquad y>0. \tag{1.7}$$

The mean, μ, and variance, σ^2, of the beta distribution are given by [6]

$$\mu = \frac{a}{a+b} \tag{1.8}$$

$$\sigma^2 = \frac{ab}{(a+b)^2(a+b+1)}, \tag{1.9}$$

or

$$\sigma^2 = r\mu(1-\mu), \tag{1.10}$$

where

$$r \equiv (a+b+1)^{-1}. \tag{1.11}$$

Once the conditional distribution and the prior distribution are chosen, the marginal distribution can then be evaluated. Upon substitution of Eqs. (1.1) and (1.5) as the conditional distribution and the prior distribution, respectively, into Eq. (1.3), the marginal distribution becomes

$$\begin{aligned} h(k;n,a,b) &= \int_0^1 f(k|p;n)\, g(p;a,b)\, dp \\ &= \frac{\binom{n}{k}}{B(a,b)} \int_0^1 p^{a+k-1}\,(1-p)^{b+n-k-1}\, dp \\ &= \binom{n}{k} \frac{B(a+k,b+n-k)}{B(a,b)}, \quad k=0,1,2,\ldots,n\ . \end{aligned} \tag{1.12}$$

Equation (1.12) gives the probability of obtaining exactly k failures in n tries for a component randomly selected from a sample whose distribution of failure probabilities is described by the beta distribution. This particular marginal distribution is called the "beta-binomial" or "hyperbinomial" whose mean and variance are given by [3]

$$E(k) = n\mu, \tag{1.13}$$

$$V(k) = n\mu(1-\mu)\ [1+r(n-1)], \tag{1.14}$$

where μ and r are defined in Eqs. (1.8) and (1.11).

Finally, the posterior distribution can be obtained by substituting the appropriate distributions from Eqs. (1.1), (1.5) and (1.12) respectively into Eq. (1.4) to give

$$\xi(p|k;n,a,b) = \frac{f(k|p;n)g(p;a,b)}{h(k;n,a,b)}$$

$$= \frac{p^{a+k-1}(1-p)^{b+n-k-1}}{B(a+k,b+n-k)} \quad , \quad 0<p<1 \ . \qquad (1.15)$$

This posterior distribution is also a beta distribution with

$$E(p) = \frac{a+k}{a+b+n} \qquad (1.16)$$

and

$$\mathrm{Var}(p) = \frac{(a+k)(b+n-k)}{(a+b+n)^2(a+b+n+1)} \ . \qquad (1.17)$$

Often E(p) from Eq. (1.16) is used as an estimator of the failure probability for a component experiencing k failures in n starts and which belongs to a class whose failure probabilities are distributed according to a beta function with parameters a and b. It can be seen from Eqs. (1.8) and (1.16) that after the actual data are observed, the estimator of a failure probability, k/n, is revised by using the mean of posterior distribution, i.e., one simply adds the number of failures to the numerator and number of tries to the denominator of Eq. (1.8). This advantage of the prior conjugate concept makes successive uses of Eq. (1.4) simple.

It should be further noted that, while posterior is also a beta distribution, it has larger parameters (a+k and b+n-k) than does the beta prior. These larger parameters produce a smaller variance which corresponds to more knowledge or more certainty about p. This is intuitively reasonable since the description of p is based on both prior information and actual experimental knowledge. Consequently, one would expect a higher degree of certainty about p for this case than the case in which only prior information or only actual experimental knowledge is used.

The posterior distribution of the failure probability describes a particular component which has experienced k failures in n attempts and whose prior distribution is assumed to be the beta distribution with parameters a and b. The prior, on the other hand, describes the variation or uncertainty in the failure probability for that component before the new information is observed. Therefore, before observing sample data, the failure probability, p, may be estimated by the mean of the prior, while after the sample data observed, the failure probability, p, is calculated from the mean of the posterior.

1.3 Construction of Prior Distribution

As previously stated, any probability distribution over the interval [0,1] can serve as a prior model. There are no rigid rules for selecting a particular prior model. Some analysts may simply choose the prior model with specified parameters which they believe to be the best representation of their knowledge about the failure probabilities of concern. If an analyst is willing to use the beta prior model, parameters a and b of the beta distribution can be determined by the method proposed by Martz and Waller [8]. This method uses expert opinion to quantify some aspects of the prior distribution, and from this quantification, values for the prior parameters are then deduced. Suppose an engineer can estimate the mean, 95-percentile and

5-percentile of the prior distribution, then any two of these values are sufficient to specify values for a and b. However, when estimators for three of these values are given, no unique solution for the parameters which is compatible with all the information is generally possible.

In the preceding analysis, the prior distribution was determined totally by one engineer's belief. Two analysts will not necessarily come up with the same distribution for a given problem owing to different degrees of belief concerning a failure probability. In an effort to remove some subjectivity from the analysis, empirical Bayes methods have been developed. These techniques use observed historical data to estimate either the parameters of a given prior model or to estimate the prior model itself. A brief introductory exposition on the empirical Bayes methods has been presented by Krutchkoff [9]. Suppose a sequence of N attribute test data have been measured for the same or similar components. Further, suppose an analyst is not willing to assume any particular functional form of the prior distribution. With the assumption that each set of attribute test data is obtained independently and can be described by the same distribution, these data can be used to estimate the prior distribution. Several techniques to estimate the prior distribution soley from observed data have been proposed by Copas [10], Griffin and Krutchkoff [11], Lemon and Krutchkoff [12] and Martz and Lian [13]. The simplest method, suggested by Lemon [14], approximates the unknown and unspecified prior model by a discrete probability model (or frequency distribution) of the data.

As an alternative to a purely empirical approach, an engineer may decide that the beta prior model is sufficiently flexible in shape and can thus adequately reflect his prior beliefs about the failure probability distribution for a particular class of components. The only problem remaining is to estimate the parameters a and b of the beta distribution from previously observed attribute data. In essence, this intermediate approach assumes a functional form for the prior distribution but bases the values for the parameters solely on empirical failure data. In this way the often contradictory expert opinion is no longer used. However, with this approach there is no known "best" method to obtain values for the prior parameters from the observed failure data. Several methods are possible each of which can yield different values for the prior parameters for the same observed failure data.

1.4 Scope of Study

It is the objective of this study to determine the "best" techniques for estimating the parameters a and b of the prior beta distribution from observed failure data. Shultis and Eckhoff [15] investigated three parameter estimation techniques, namely, the prior matching moments, the weighted marginal matching moments, and the marginal maximum likelihood method. These techniques were used to analyze failure data obtained from standby diesel generators at many U.S. nuclear power plants. The results of this study shows that the different estimation techniques often produce quite different values for the parameters of the prior distribution.

In this study the properties of these prior parameter estimation techniques are investigated in detail. To determine which parameter estimation technique is the most conservative or yields parameter estimators closest to the true values, it is necessary to determine the distribution of the parameter estimators for each estimation method. Two additional estimation methods, namely, the prior maximum likelihood method and a weighted marginal matching moments method are also investigated in this study. The details of these five methods will be presented in the next chapter.

For this investigation, multiple sets of failure data in various sample sizes (pairs of failure data - number of failures and number of attempts) were generated randomly from a known beta-binomial distribution. With these simulated failure data, the distribution of the prior parameter estimates are determined empirically for each estimation technique, and from these distributions many properties of the five estimation techniques are then determined. In particular, the biasedness, mean-squared error and median of each estimation technique are examined. The distribution of the mean and variance of a failure probability are also investigated. Also presented are the distribution of 95-percentiles of the failure probability and the distribution of the probability that a failure probability is greater than the true 95-percentile. Finally, both exact and approximate methods for obtaining lower bounds of the variances and covariance of the parameter estimators are computed. These bounds, which are based on the Cramer Rao-Frechet inequality for a covariance matrix, are compared to the values of the variances and covariance obtained from the simulated data.

2.0 PRIOR PARAMETER ESTIMATION TECHNIQUES

Before a Bayesian analysis or compound classical analysis can be performed for a set of components, the prior distribution for the set must first be determined. In this study, the analysis of failure-on-demand attribute data is considered, and thus, because of the advantages afforded by using a prior conjugate to the likelihood distribution (as is discussed in the previous chapter), a beta distribution is assumed for the prior distribution. Moreover, this family can assume a wide variety of shapes for different values of its two parameters, and thus can approximate almost any realistic prior distribution.

The parameters "a" and "b" of the beta distribution in Eq. (1.5) must be given definite values in order to specify uniquely a particular prior distribution. The problem, therefore, is to determine point estimators of the prior distribution parameters appropriate to a given class of components. In general this parameter estimation is done by (i) intuition, (ii) past experience, or (iii) from a fit to experimental failure data from similar components or past test data from a single component.

In using methods (i) and (ii), an experimenter chooses the values of the beta parameters so that the beta distribution takes a form which he believes best represents his prior knowledge about the failure probability "p" of a class of component. If on the basis of engineering judgement and experience, he believes that, for example $P(p \geq 0.13) = 0.05$, the prior distribution should exhibit this property. This is a subjective step in the analysis since the values of parameters a and b depend solely upon an experimenter's belief. In an effort to remove some of this subjectivity and to be more compatible with a compound classical analysis, method (iii) is used in this study to estimate the prior parameters based only upon observed failure data. Such a technique which uses only observed historical data are commonly referred to as an empirical Bayes' method or a compound classical analysis.

For any particular plant component, given only its number of failures and total number of starts, the data are not sufficient to estimate a and b. However, if one believes several sets of data belong to the same class and consequently have the same prior distribution, then the observed data can be matched to the assumed form of the prior distribution. Generally, there is no unique way to fit the prior distribution to the observed historical data. In the following sections, five empirical methods are presented.

2.1 Method of Matching Moments of the Prior Distribution to Data Moments (PMM)

One method for determining point estimators is to equate sample moments found from the observed data to the corresponding expressions for the prior model involving the distribution parameters. From a statistical point of view, a moment is defined as the average deviation of a set of data about a point. Usually, moments about the mean are of interest (the mean itself is a first moment about the origin, i.e.,

$$\mu \equiv E(X)$$

where X is the random variable).

Formally, the k-th moment about the mean (μ) of a random variable X is defined as $E(X-\mu)^k$, which can be determined by using the usual rules of expectation as

$$E(X-\mu)^k \equiv \begin{cases} \sum (X^{k}-\mu)^k \, P(X=x) & \text{if } x \text{ is discrete} \\ \int_{-\infty}^{\infty} (X-\mu)^k \, f(X) \, dX & \text{if } x \text{ is continuous.} \end{cases}$$

These moments are measures used to describe characteristics of the distribution of X; for example, a mean (i.e., a first moment about the origin) is a measure of the central location, a variance (i.e., a second moment about a mean) is a measure of dispersion, a third moment about a mean is a measure of the skewness.

In case of a beta distribution, the mean and variance are given by (Eqs. (1.8) and (1.9))

$$\mu = \frac{a}{a+b} \tag{2.1a}$$

$$\sigma^2 = \frac{ab}{(a+b)^2(a+b+1)} \, . \tag{2.1b}$$

Suppose from past data of N components, k failures out of n tries for the i-th component were observed. By the assumption that these data belong to the same population and are hence described by the same beta distribution, the observed mean and variance (which are the estimators of mean and variance of the beta distribution) can be calculated from the sample data by the following expressions:

$$\hat{\mu}_{ob} = \frac{1}{N} \sum_{i=1}^{N} p_i \, , \tag{2.2a}$$

$$\hat{\sigma}^2_{ob} = \frac{1}{N-1} \sum_{i=1}^{N} (p_i - \hat{\mu}_{ob})^2, \tag{2.2b}$$

where p_i is the failure probability of i-th component and is estimated by the proportion k_i/n_i.

Since $\hat{\mu}_{ob}$ and $\hat{\sigma}^2_{ob}$ are the estimators of μ and σ^2 respectively, it is not unreasonable to match the observed values, which depend only the observed data (Eqs. (2.2a) and (2.2b)), to the expressions for the theoretical mean and variance of the beta distribution (Eqs. (2.1a) and (2.1b)) as follows:

$$\hat{\mu}_{ob} = \mu \equiv \frac{a}{a+b} \quad , \tag{2.3a}$$

and

$$\hat{\sigma}^2_{ob} = \sigma^2 \equiv \frac{ab}{(a+b)^2(a+b+1)} . \tag{2.3b}$$

The above relations can be solved for a and b in terms of $\hat{\mu}_{ob}$ and $\hat{\sigma}^2_{ob}$ to give

$$a = \frac{\hat{\mu}^2_{ob}}{\hat{\sigma}^2_{ob}} (1-\hat{\mu}_{ob}) - \hat{\mu}_{ob} \tag{2.4a}$$

$$b = \frac{\hat{\mu}_{ob}}{\hat{\sigma}^2_{ob}} (1-\hat{\mu}_{ob}) + \hat{\mu}_{ob} - 1. \tag{2.4b}$$

One disadvantage of this method is that component failure probability is not actually observed, but is estimated by the ratio of the number of failures to the number of tries. This estimation of b may appear to introduce a questionable approximation especially for low probability events with a small number of tries, n_i, for which one is likely to observe zero failures. For the beta distribution, the major advantage of this moment method is its simplicity and the existence of a closed-form solution for the parameter estimators. However, these solutions for the parameter estimators do not necessarily yield positive values as are required for the beta parameters unless a (usually unrestrictive) condition is satisfied.

To derive this limitation, rewrite Eq. (2.4a) for a as

$$a = \hat{\mu}_{ob}\left[\frac{\hat{\mu}_{ob}}{\hat{\sigma}^2_{ob}}(1-\hat{\mu}_{ob}) -1\right]. \tag{2.5}$$

Obviously the term $(1-\hat{\mu}_{ob})$ is always positive. Therefore, Eq. (2.5) yields positive value of a if and only if

$$\frac{\hat{\mu}_{ob}}{\hat{\sigma}^2_{ob}}(1-\hat{\mu}_{ob}) > 1,$$

or

$$\hat{\mu}_{ob}(1-\hat{\mu}_{ob}) > \hat{\sigma}^2_{ob} . \tag{2.6}$$

Finally, if the estimate for a is positive, then so must be the estimate of b since from Eq. (2.3a)

$$b = \frac{a(1-\hat{\mu}_{ob})}{\hat{\mu}_{ob}} > 0 \qquad \text{if} \quad a>0. \tag{2.7}$$

Nevertheless, for low failure probability case, it can be shown that this estimation technique almost always give positive and hence realistic values for the parameter estimates by rewriting Eq. (2.4a) as

$$a = \frac{\hat{\mu}_{ob}}{\hat{\sigma}^2_{ob}}[\hat{\mu}_{ob} - \hat{\mu}^2_{ob} - \hat{\sigma}^2_{ob}].$$

Upon substitution for $\hat{\mu}_{ob}$ and $\hat{\sigma}^2_{ob}$ from Eqs. (2.2a) and (2.2b) the above relation becomes

$$a = \frac{\hat{\mu}_{ob}}{\hat{\sigma}^2_{ob}} \left(\frac{1}{N} \sum_i p_i - \frac{1}{N-1} \sum_i p_i^2 + \frac{1}{(N-1)N^2} \left(\sum_i p_i \right)^2 \right)$$

$$a = \frac{\hat{\mu}_{ob}}{\hat{\sigma}^2_{ob}} \cdot \frac{1}{N} \left(\sum_i p_i \; (1 - \frac{N}{N-1} p_i) \right) + \frac{\hat{\mu}_{ob}}{\hat{\sigma}^2_{ob}(N-1)N^2} (\sum_i p_i)^2$$

$$a > \frac{1}{N} \cdot \frac{\hat{\mu}_{ob}}{\hat{\sigma}^2_{ob}} [\sum_i p_i \; (1 - \frac{N}{N-1} p_i)].$$

The right hand side of this inequality would be positive, if $0 \leq p_i \leq \frac{N-1}{N}$. For the most restrictive case (N=2), positive estimates of a are always obtained if $0 \leq p_i \leq \frac{1}{2}$ which is very likely to occur for low failure probability data. As the sample size (N) becomes larger, the upper limit on p_i becomes even higher. For sufficiently large N, the restriction on the p_i value vanishes. Furthermore, Eq. (2.7) shows that estimates of b always take the same sign as that of a.

Finally, it should be mentioned that this restriction for realistic estimates arises because of the definition of the sample variance. If the observed variance of Eq. (2.2b) had been defined with the factor 1/N rather than the usual 1/(N-1), then the same analysis as presented about would show that estimates for a and b given by Eqs. (2.4) would always be non-negative.

2.2 Maximum Likelihood Method Based on the Prior Distribution (PML)

The maximum likelihood method is one of the most popular methods to estimate parameters of probability functions from observed sample data. For many distributions the method of maximum likelihood, developed by Fisher in the 1920's, gives about as accurate estimates as are possible from a given set of data [16].

The method of maximum likelihood estimates the value of a distribution parameter by selecting the most likely sample space from which a given sample could have been drawn. In other words, the sample space is selected which would yield the observed sample more frequently than any other sample space. The distribution parameter corresponding to this space is called the maximum likelihood estimator.

Suppose that a random variable X has a distribution which is described by the known probability density function f(X;c) with unknown parameter c. A sample of N independent observations is drawn, producing the set of values $(x_1,x_2,\ldots,x_N)$. A likelihood function of this particular sample is defined by the relation

$$L(x_1,x_2,\ldots,x_N;c) \equiv \prod_{i=1}^{N} f(x_i;c). \qquad (2.8)$$

Equation (2.8) is simply the joint probability of obtaining $x_1,x_2,\ldots,x_N$ in the same sample set provided that each sample is drawn independently. The maximum likelihood estimator $\hat{c}$ is the value of c which maximizes the likelihood function. Mathematically, $\hat{c}$ is the solution of

$$\frac{\partial L}{\partial c}(x_1,x_2,\ldots,x_N;c) = 0 . \qquad (2.9)$$

In many cases it is more convenient to maximize the logarithm of L rather than L itself. This transformation will not change the optimal solution, but very often will make finding the derivative easier. The values of c at which L and $ln L$ have extrema are identical. Thus $\hat{c}$ also can be obtained from solution of

$$\frac{\partial \, ln L}{\partial c}(x_1, x_2, \ldots, x_N; c) = 0 \ . \tag{2.10}$$

For m parameters, the maximum likelihood estimators are the solution of m simultaneous equations found by setting to zero the first derivatives of logarithm of likelihood function with respect to each parameter, i.e.,

$$\frac{\partial \, ln L}{\partial c_i} = 0, \qquad i=1,2,\ldots,m \ . \tag{2.11}$$

The maximum likelihood estimator has several desirable properties other than the intuitively appealing property that it maximizes the likelihood function. Among the most important features of this method is the behavior of the likelihood function as the number of data points (N) becomes large (the so-called asymptotic properties). For large N, the estimators from this method possess many of the desirable properties of a good estimator, i.e., unbiased, minimum variance, most efficient and consistent [17]. Furthermore, the likelihood function approaches a normal distribution with mean $\hat{c}$ and variance $1/[E\{(\frac{\partial \, ln L}{\partial c})^2\ ]$ which is the Cramer-Rao lower bound for any unbiased variance estimator determined by any method [17].

However, for small sample sizes, the maximum likelihood estimator is not always unbiased and many of its other desirable asymptotic properties are lacking. Moreover, it frequently happens that the likelihood equation (Eq. (2.11)) is not solvable analytically. In such cases numerical or graphical methods must be used to obtain the maximum likelihood estimators.

In this section, the maximum likelihood method is used to obtain estimates of the prior parameters by constructing a likelihood function based on the beta distribution (or the prior) as given by

$$L(p_1,p_2,\ldots,p_N;a,b) \equiv \prod_{i=1}^{N} g(p_i;a,b) \qquad (2.12)$$

where $g(p_i;a,b)$ is the beta distribution from Eq. (1.5).

This likelihood function is the probability of observing $p_1,p_2,\ldots,$ p_n as values for the failure probabilities from components $1,2,\ldots,N$ respectively with the assumption that each component is observed independently. As mentioned in Section 2.1, the component failure probability is not directly observed for the failure-on-demand problem considered in this study but is estimated by k_i/n_i (which is the maximum likelihood estimator of p_i). The maximum likelihood estimators $\hat{a}$ and $\hat{b}$ are calculated in such a way that the probability of obtaining the observed values of k_i/n_i is maximized, i.e., they are chosen as the values of a and b which maximize the likelihood function (Eq. (2.12). Thus, mathematically, the estimators $\hat{a}$ and $\hat{b}$ are the solution to

$$\frac{\partial}{\partial a}\ln L(a,b) = 0 \qquad (2.13)$$

$$\frac{\partial}{\partial b}\ln L(a,b) = 0.$$

Upon substitution of the explicit form of the beta prior function, $g(p_i;a,b)$ from Eq. (1.5), the likelihood equations (Eq. (2.13)) become

$$\psi(a) - \psi(a+b) - N^{-1} \sum_{i=1}^{N} \ln p_i = 0$$

$$\psi(b) - \psi(a+b) - N^{-1} \sum_{i=1}^{N} \ln(1-p_i) = 0 \qquad (2.14)$$

where $\psi(z) \equiv d\ [\ln\Gamma(z)]/dz$ is the digamma function.

The solution to these simultaneous equations cannot be obtained analytically; however, if $\hat{a}$ and $\hat{b}$ are not too small the following approximate results may be used [6].

$$\hat{a} \simeq 1/2\ [1 - \prod_{i=1}^{N} (1-p_i)^{1/N}][1 - \prod_{i=1}^{N} p_i^{1/N} - \prod_{i=1}^{N} (1-p_i)^{1/N}]^{-1}$$

$$\hat{b} \simeq 1/2\ [1 - \prod_{i=1}^{N} p_i^{1/N}][1 - \prod_{i=1}^{N} p_i^{1/N} - \prod_{i=1}^{N} (1-p_i)^{1/N}]^{-1}. \qquad (2.15)$$

This approximate solution may also be used as starting values for an iterative numerical solution of the likelihood equations.

One major disadvantage of the maximum likelihood method based on the beta prior distribution is that when the number of failures k is observed to be zero, the failure probability of that component is also estimated to be zero, and the $\ln p_i$ term used in Eq. (2.14) becomes singular. Since it is quite probable to observe zero failure for a small number of tries, this method is not suitable for the analysis of data from components with expected low failure probabilities. Consequently, little use was made of this technique in this analysis.

2.3 Method of Weighted Marginal Matching Moments (WMMM) [3]

The bias of this method is the same as that of PMM (i.e., matching moments). In the PMM method, the effect of different values of n_i is ignored, because only the ratio of k_i/n_i is used to calculate the sample statistics. From the fact that the variance of k with a larger n is smaller than that with a smaller n, the sample with the larger n should intuitively have more value in computing sample statistics used in the parameter estimation. To do so, a weighting scheme which is a function of n is incorporated into the estimation procedure to favor samples with large n. This concept was originally proposed by Joel C. Kleinman [3].

In this section, the weighted marginal matching moments theory is reviewed. Kleinman proposed the following statistics: (13)

$$\hat{p} = \frac{1}{w} \sum_{i=1}^{N} w_i \frac{k_i}{n_i} , \tag{2.16}$$

$$S = \sum_{i=1}^{N} w_i \{\hat{p} - \frac{k_i}{n_i}\}^2 , \tag{2.17}$$

where $w = \sum_{i=1}^{N} w_i$, and w_i is the weight assigned to the i-th sample for this derivation (w_i is quite arbitrary). Recall from Eqs. (1.13) and (1.14), the mean and variance of a beta-binomial distribution are given by

$$E(k) = n\mu , \tag{2.18a}$$

$$\mathrm{Var}(k) = n\mu(1-\mu)\ [1 + r(n-1)], \qquad (2.18b)$$

where

$$\mu = a\ (a+b) \quad \text{and} \quad r = (a+b+1)^{-1}.$$

Then rewriting Eqs. (2.18a) and (2.18b) in term of k/n, which is the estimator of p, $\hat{p}$, one obtains

$$E(k/n) = \mu \qquad (2.19a)$$

$$\mathrm{Var}(k/n) = \frac{1}{n} \cdot \mu(1-\mu)\ [1+r(n-1)]. \qquad (2.19b)$$

Further, expectations of the statistics from Eqs. (2.16) and (2.17) are given by Eqs. (2.19a) and (2.20) [see Appendix A],

$$E(S) = \mu(1-\mu) \sum_{i=1}^{N} \frac{w_i}{n_i} (1 - \frac{w_i}{w}) + r\mu(1-\mu)\ [\sum_{i=1}^{N} w_i\ (1 - \frac{w_i}{w}) - \sum_{i=1}^{N} \frac{w_i}{n_i} (1 - \frac{w_i}{w})]\ . \qquad (2.20)$$

By setting the statistics in Eqs. (2.16) and (2.17) equal to their expected values, (Eqs. (2.19a) and (2.20) respectively), the following relations are obtained:

$$\hat{p} = \mu \qquad (2.21)$$

and

$$S = E(S). \qquad (2.22)$$

Substituting for E(S) on the right hand side of Eq. (2.22) by Eq. (2.20) and rearranging, one obtains

$$\hat{r} = \frac{S - \hat{p}\hat{q}\,[\sum_{i=1}^{N} \frac{w_i}{n_i}(1 - \frac{w_i}{w})]}{\hat{p}\hat{q}\,[\sum_{i=1}^{N} w_i (1 - \frac{w_i}{w}) - \sum_{i=1}^{N} \frac{w_i}{n_i}(1 - \frac{w_i}{w})]}, \qquad (2.23)$$

where $\hat{q} = 1-\hat{p}$.

The estimates for the prior mean and variance are given by Eq. (2.22) and

$$\hat{\sigma}^2 = \hat{\mu}(1-\hat{\mu})\,\hat{r}\ . \qquad (2.24)$$

Upon substitution of $\hat{r}$ from Eq. (2.23), the estimate of variance becomes

$$\hat{\sigma}^2 = \hat{\mu}(1-\hat{\mu})\frac{S - \hat{p}\hat{q}\,[\sum_{i=1}^{N} \frac{w_i}{n_i}(1 - \frac{w_i}{w})]}{\hat{p}\hat{q}\,[\sum_{i=1}^{N} w_i (1 - \frac{w_i}{w}) - \sum_{i=1}^{N} \frac{w_i}{n_i}(1 - \frac{w_i}{w})]}. \qquad (2.25)$$

Once $\hat{\mu}$ and $\hat{\sigma}^2$ of the prior distribution are calculated from Eqs. (2.21) and (2.25), the parameter estimates $\hat{a}$ and $\hat{b}$ are found by the same way as in Section 2.1.

Kleinman chose the weight w_i, used to calculate statistics $\hat{p}$ and S in Eqs. (2.16) and (2.17), to be the inverse of variance of p_i $(p_i=k_i/n_i)$ from Eq. (2.19b), namely

$$w_i = \frac{n_i}{\mu(1-\mu)[1+r(n_i-1)]}\ . \qquad (2.26)$$

Since the term $\mu(1-\mu)$ is the same for all w_i, the expression for the weight is reduced to

$$w_i = \frac{n_i}{1+r(n_i-1)}\ . \qquad (2.27)$$

To calculate $\hat{\beta}$ and $\hat{\sigma}^2$, the weight, w_i, must be known, which from Eq. (2.27) implies that r (or $\hat{\sigma}^2$), must be specified.

One possible way to obtain values of $\hat{\beta}$ and $\hat{\sigma}^2$ is to use an iteration process. The following iteration scheme was adopted in this study:

1. Choose $r=0$, therefore $w_i=n_i$ from Eq. (2.27),
2. Solve for $\hat{\beta}$ and $\hat{\sigma}^2$ from Eqs. (2.21), (2.16), (2.17) and (2.25),
3. Use values of $\hat{\beta}$ and $\hat{\sigma}^2$ from step 2 to calculate new values of r and w_i from Eqs. (2.24) and (2.27),
4. Continue iterating steps 2 and 3 until $\hat{\beta}$, $\hat{\sigma}^2$ and w_i converge to certain values.

Kleinman [3] suggests that better estimates of $\hat{\beta}$ and $\hat{\sigma}^2$ are obtained if S in Eq. (2.25) is replaced by $(N-1)S/N$. This suggestion was incorporated in the numerical procedures developed for this work. In practice the calculation procedure of this method is rather complicated and a solution for the estimate is not always obtained or is sometimes unrealistic.

2.4 Method of Unweighted Marginal Matching Moments (UWMMM)

In Kleinman's method (WMMM), components with more number of startups or tries n received more emphasis in the calculation of the parameter estimators. It might be argued that it is unjustified to pay more attention to one component over another, particularly if each component is viewed as having its own distinct, but unknown failure probability. A more reasonable approach may be to apply the same weight to every component.

Upon substitution of $w_i=1$, $i=1,2,\ldots,N$ into Eqs. (2.16), (2.17), and (2.25), one has

$$\hat{p} = \sum_{i=1}^{N} \frac{k_i}{n_i} ,$$

$$S = \sum_{i=1}^{N} \left(\frac{k_i}{n_i} - \hat{p}\right)^2 ,$$

and

$$\hat{\sigma}^2 = \hat{\mu}(1-\hat{\mu}) \frac{\frac{S}{\hat{p}\hat{q}} - \frac{N-1}{N} \sum_{i=1}^{N} \frac{1}{n_i}}{(N-1) \left[1 - \frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_i}\right]}$$

where $\hat{q} = 1-\hat{p}$. With the above relations and Eq. (2.21), estimators of $\hat{\mu}$ and $\hat{\sigma}^2$ can be obtained. Then the parameter estimators of $\hat{a}$ and $\hat{b}$ can be evaluated in the same way as was done in Section 2.1.

By using the same weight, the previous iterative scheme used with the WMMM method is unnecessary which is a major computational advantage of this method over the WMMM. Moreover, as will be shown, the UWMMM method often yields better results than does the WMMM.

2.5 Maximum Likelihood Method Based on the Marginal Distribution (MML)

In Section 2.2, parameter estimates are obtained by using the maximum likelihood technique based on the beta prior distribution. However, the problems encountered with this technique for a case of zero number of failure makes this method unsuitable to a low failure probability situation. An alternative to the technique of Section 2.2 is to define the likelihood function in terms of the beta-binomial marginal distribution. This alternate approach is also appealing because the actual observed data (i.e., number of failures k_i out of n_i tries) are used, whereas in the previous prior based maximum likelihood method of Section 2.2, the failure probabilities p_i were required (which were not actually observed) and had to be estimated as k_i/n_i.

The likelihood function is defined as

$$L(k_1,k_2,\ldots,k_N;n_1,n_2,\ldots,n_N,a,b) \equiv \prod_{i=1}^{N} h(k_i;n_i,a,b) \tag{2.28}$$

where $h(k_i;n_i,a,b)$ is the beta-binomial marginal distribution given by Eq. (1.12). This likelihood function is the probability of obtaining $k_1,k_2,\ldots k_N$ failures in $n_1,n_2,\ldots,n_N$ tries from N components each of whose failure distribution is described by the beta-binomial marginal distribution of Eq. (1.12) with parameters a and b under the assumption that data from each component are obtained independently. Substitution of Eq. (1.12) into Eq. (2.28) yields

$$L(a,b) \equiv L(k_1,\ldots,k_N;\ n_1,\ldots,n_N,a,b)$$

$$= \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right)^N \prod_{i=1}^{N} C_i \, \frac{\Gamma(a+k_i)\Gamma(b+n_i-k_i)}{\Gamma(a+b+n_i)} \, , \tag{2.29}$$

where the binomial coefficient is given by

$$C_i \equiv \binom{n_i}{k_i} = \frac{\Gamma(n_i+1)}{\Gamma(k_{i+1})\Gamma(n_i-k_i+1)} \quad .$$

The values of a and b which maximize the above likelihood function are called the marginal maximum likelihood estimates, $\hat{a}$ and $\hat{b}$. As in Section 2.2, a logarithmic transformation makes the subsequent numerical computation easier. Therefore, $\hat{a}$ and $\hat{b}$ are obtained from solutions to

$$\frac{\partial}{\partial a} \ln L(a,b) = 0$$

$$\frac{\partial}{\partial b} \ln L(a,b) = 0$$

or explicitly

$$N[\psi(a+b)-\psi(a)] + \sum_{i=1}^{N} [\psi(a+k_i)-\psi(a+b+n_i] = 0 \tag{2.30a}$$

and

$$N[\psi(a+b)-\psi(b)] + \sum_{i=1}^{N} [\psi(b+n_i-k_i)-\psi(a+b+n_i)] = 0 \tag{2.30b}$$

where $\psi(z) \equiv \frac{d}{dz} [\ln\Gamma(z)]$, the digamma function. The numerical solution of these two simultaneous is obtained by standard numerical techniques [such as the Newton-Raphson method, with the PMM estimators (Section 2.1) as the starting values].

This method, unlike the PML method (Section 2.2), is seen to present no special difficulties for the case of zero number of failures. However, it is found that a solution from this method is not always obtained since for some samples the maximum point of the likelihood function is $a=b=\infty$. Also, the MML estimators, when they exist, occasionally are much too large to be accepted as reasonable. Finally it should be pointed out that of the five methods investigated in this study, the maximum likelihood methods are the most complicated and, as will be shown in the next chapter, they do not always yield the best results for small sample sizes.

2.6 Lower Bounds on the Variances of Prior Parameter Estimates [16,17]

One of the most attractive features of the maximum likelihood method is that, besides yielding estimates of the parameters, this method also yields lower bounds for the variances and the covariance of the parameters. These lower bounds can often be used as useful approximations to the variances and covariance of the estimates.

For N independent observations, $x_1, x_2, \ldots, x_N$, where each observation is from a distribution $f(x;\underline{c})$, the likelihood function is defined by

$$L(x_1, x_2, \ldots, x_N; \underline{c}) \equiv \prod_{i=1}^{N} f(x_i;\underline{c}) \tag{2.31}$$

where x and $\underline{c}$ represent the sample random variable and parameter vector of size M, respectively. The maximum likelihood estimators of $\underline{c}$ are denoted by $\underline{\hat{c}}$ which are those values of the parameters which maximize L, i.e.,

$$\frac{\partial}{\partial c_i} L(x_1, x_2, \ldots, x_N; \underline{c}) \Big|_{c_i = \hat{c}_i} = 0, \quad i=1,2,\ldots,M \tag{2.32}$$

or equivalently maximize $\ell n L$, i.e.,

$$\frac{\partial}{\partial c_i} \ell n L(x_1, x_2, \ldots, x_N; \underline{c}) \Big|_{c_i = \hat{c}_i} = 0, \quad i=1,2,\ldots,M \quad . \tag{2.33}$$

The elements of the information matrix [I] are defined [10] as

$$I_{ij}(\underline{c}) \equiv E[\{\frac{\partial \ell}{\partial c_i}\}\{\frac{\partial \ell}{\partial c_j}\}], \quad i,j=1,2,\ldots,M \tag{2.34}$$

where $\ell \equiv \ell n L(x_1, x_2, \ldots, x_N; \underline{c})$, and the expectation is with respect to the likelihood function L. These elements of the information matrix may be written in alternative form as [17]

$$I_{ij}(\underline{c}) = -E[\frac{\partial^2 \ell}{\partial c_i \partial c_j}] \;,\; i,j=1,2,\ldots,M \qquad (2.35)$$

By definition of expectation, the elements of the information matrix $I_{ij}(\underline{c})$ can thus be computed from

$$I_{ij}(\underline{c}) = -E\left(\frac{\partial^2 \ell}{\partial c_i \partial c_j}\right) = -\int dx_1 \int dx_2 --- \int dx_N \frac{\partial^2 \ell}{\partial c_i \partial c_j} L(\underline{c}|x_1,x_2,\ldots,x_N),$$

$$i,j=1,2,\ldots,M \qquad (2.36)$$

where the integration (or summation in the case of a discrete distribution) is over all possible values of variables $x_1,x_2,\ldots,x_N$. If one assumes that the distribution of the likelihood function with respect to each parameter is symmetrical, then [17]

$$E\left(\frac{\partial^2 \ell}{\partial c_i \partial c_j}\right) \approx \left(\frac{\partial^2 \ell}{\partial c_i \partial c_j}\right)\Bigg|_{\underline{c} = \underline{\hat{c}}} . \qquad (2.37)$$

One of the most important theorems about the maximum likelihood method is known as the Cramer-Rao-Frechet inequality [17] which states that if $\underline{\hat{c}}$ are any unbiased estimators of $\underline{c}$, then

$$[\sigma] \geq [A] \;, \qquad (2.38)$$

where elements σ_{ii} and σ_{ij}, $i \neq j$; $i,j=1,2,\ldots M$ of matrix $[\sigma]$ are variances of parameters c_i and covariances of parameter c_i and c_j respectively, and matrix [A] is the inverse of the information matrix.

In other words, lower bounds for any unbiased variances and covariance of the parameter estimators $\underline{\hat{c}}$ can be obtained from the elements of the inverse of the information matrix as

$$\left.\begin{aligned} &\text{var}(\hat{c}_i) = \sigma_{ii} \leq A_{ii}, \\ \text{and} \quad &\\ &\text{cov}(\hat{c}_i,\hat{c}_j) = \sigma_{ij} \leq A_{ij}. \end{aligned}\right\} \qquad (2.39)$$

As stated in the previous section, among the most important features of the method of maximum likelihood is the behavior of the likelihood function as the number of data points, N, becomes large. For large N, the likelihood function approaches a normal distribution with mean $\underline{c}$ and variances

$$\sigma_{ii} = A_{ii}$$

Asymptotic properties of the likelihood function guarantee that

$$\lim_{N\to\infty} E\,[\hat{c}_i] = c_i \,, \tag{2.40}$$

$$\lim_{N\to\infty} \mathrm{var}[\hat{c}_i] = A_{ii} \,, \tag{2.41}$$

and

$$\lim_{N\to\infty} \mathrm{cov}[c_i,c_j] = A_{ij} \,. \tag{2.42}$$

In addition, the approximation of Eq. (2.37) is valid when N is sufficiently large.

To apply the above results to the problem of estimating the variances and covariance of the two parameters of the prior beta distribution, the information matrix is constructed for Eq. (2.28) which is the likelihood function based on the beta-binomial distribution, as

$$[I(a,b)] = -\begin{pmatrix} E[\frac{\partial^2 \ell}{\partial a^2}] & E[\frac{\partial^2 \ell}{\partial a \partial b}] \\ E[\frac{\partial^2 \ell}{\partial a \partial b}] & E[\frac{\partial^2 \ell}{\partial b^2}] \end{pmatrix} , \tag{2.43}$$

where $\ell \equiv \mathit{ln} L(a,b)$. The second derivatives of the logarithm of the likelihood function are given by

$$\frac{\partial^2 \ell}{\partial a^2} = N\,\{\psi'(a+b)-\psi'(a)\} + \sum_{i=1}^{N} \{\psi'(a+k_i) - \psi'(a+b+n_i)\} \tag{2.44}$$

$$\frac{\partial^2 \ell}{\partial b^2} = N\ \{\psi'(a+b)-\psi'(b)\} + \sum_{i=1}^{N} \{\psi'(b+n_i-k_i)-\psi'(a+b+n_i)\} \tag{2.45}$$

$$\frac{\partial^2 \ell}{\partial a \partial b} = N\psi'(a+b) - \sum_{i=1}^{N} \psi'(a+b+n_i) \tag{2.46}$$

where $\psi'(Z) \equiv d^2[\ell n\Gamma(Z)]/dZ^2$ is the trigamma function.

For the failure-on-demand model of this work, the likelihood function is a discrete distribution, and thus the expectation values for the matrix elements of Eq. (2.43) are calculated by

$$E[y] = \sum_{k_1=0}^{n_1} \sum_{k_2=0}^{n_2} \cdots \sum_{k_n=0}^{n_N} y\ L(a,b), \tag{2.47}$$

where $L(a,b) \equiv \prod_{i=1}^{N} h(k_i;n_i,a,b)$.

Since

$$\sum_{k_i=0}^{n_i} h(k_i;n_i,a,b) = 1,$$

the substitution of the explicit form of the likelihood function and subsequent simplification gives the following results for the elements of the information matrix:

$$E\left(\frac{\partial^2 \ell}{\partial a^2}\right) = N\{\psi'(a+b)-\psi'(a)\} + \sum_{i=1}^{N} \sum_{k_i=0}^{n_i} \psi'(a+k_i)h(k_i;n_i,a,b) - \sum_{i=1}^{N} \psi'(a+b+n_i) \tag{2.48}$$

$$E\left(\frac{\partial^2 \ell}{\partial b^2}\right) = N\{\psi'(a+b)-\psi'(b)\} + \sum_{i=1}^{N} \sum_{k_i=0}^{n_i} \psi'(b+n_i-k_i)h(k_i;n_i,a,b) - \sum_{i=1}^{N} \psi'(a+b+n_i) \tag{2.49}$$

$$E\left(\frac{\partial^2 \ell}{\partial a \partial b}\right) = N\psi'(a+b) - \sum_{i=1}^{N} \psi'(a+b+n_i) \ . \tag{2.50}$$

The numerical evaluation of the expected values of the elements of the information matrix can be quite time consuming expecially when N and the n_i are large. Application of Eq. (2.37) allows a much more expedient, but approximate, evaluation of these matrix elements. Specifically one has

$$E\left(\frac{\partial^2 \ell}{\partial a^2}\right) \simeq \left(\frac{\partial^2 \ell}{\partial a^2}\right)_{\substack{a=\hat{a}\\ b=\hat{b}}} = N\psi'(\hat{a}+\hat{b}) - N\psi'(\hat{a}) + \sum_{i=1}^{N} \{\psi'(\hat{a}+k_i)-\psi'(\hat{a}+\hat{b}+n_i)\}, \tag{2.51}$$

$$E\left(\frac{\partial^2 \ell}{\partial b^2}\right) \simeq \left(\frac{\partial^2 \ell}{\partial b^2}\right)_{\substack{a=\hat{a}\\ b=\hat{b}}} = N\psi'(\hat{a}+\hat{b}) - N\psi'(\hat{b}) + \sum_{i=1}^{N} \{\psi'(\hat{b}+n_i-k_i)-\psi'(\hat{a}+\hat{b}+n_i)\}, \tag{2.52}$$

$$E\left(\frac{\partial^2 \ell}{\partial a \partial b}\right) \simeq \left(\frac{\partial^2 \ell}{\partial a \partial b}\right)_{\substack{a=\hat{a}\\ b=\hat{b}}} = N\psi'(\hat{a}+\hat{b}) - \sum_{i=1}^{N} \psi'(\hat{a}+\hat{b}+n_i). \tag{2.53}$$

Finally, from Eqs. (2.39) one has the following approximations for the variances and covariance of the maximum likelihood estimators:

$$\mathrm{var}(\hat{a}) \simeq [I^{-1}(\hat{a},\hat{b})]_{11}, \tag{2.54}$$

$$\mathrm{var}(\hat{b}) \simeq [I^{-1}(\hat{a},\hat{b})]_{22}, \tag{2.55}$$

and

$$\mathrm{cov}(\hat{a},\hat{b}) \simeq [I^{-1}(\hat{a},\hat{b})]_{12}. \tag{2.56}$$

3.0 SIMULATION STUDY OF PRIOR ESTIMATION TECHNIQUES

Five methods for estimation of the prior parameters were presented in the previous chapter. Previous results of using three of these methods with actual failure records of standby diesel engines from some nuclear power plants exhibited significant differences in the estimators [15]. However, it is not reasonable to conclude which of the five methods yields the best estimators from the few results obtained. To determine many important properties of the various estimators, their distributions are needed.

To construct such distributions, multiple sets of attribute failure data are necessary. With data being randomly chosen from a known beta-binomial distribution, the distribution of the prior parameter estimators could be determined for each estimation technique. From these distributions of the estimators many properties of the five estimation techniques can then be investigated. To have some criteria to determine the "best" estimation method, some desirable properties of a good estimator will be reviewed in the next section. The simulation technique for the generation of failure data will then be presented in Section 3.2.

3.1 Desirable Properties of an Estimator [18]

A major objective of this study is to determine the best method of the five estimation techniques discussed in the previous chapter. Good estimators should possess all or most of the following properties: unbiasedness, consistency, efficiency and sufficiency.

3.1.1 Unbiasedness

Assume that a sample statistic G is used as an estimator of a parameter θ. The statistic G is then said to be an "unbiased estimator" of θ if

$$E(G) = \theta \qquad (3.1)$$

where the expectation is with respect to a joint density function of random variables used to calculate the statistic G.

In other words, suppose that random samples are taken repeatedly and the value of G is evaluated for each sample. Then in the long run, if the average value of G is θ, the estimator G has the property of unbiasedness.

3.1.2 Consistency

An intuitively attractive property for an estimator to possess is that the sample estimate should have a higher probability of being close to the population value θ for a larger sample size. A statistic that has this property is called a "consistent estimator". More formally, the statistic G is a consistent estimator of θ if for any arbitrary ε,

$$P(|G-\theta| < \varepsilon) \longrightarrow 1 \quad \text{as } n\to\infty \qquad (3.2)$$

where n is the sample size.

An important distinction between unbiasedness and consistency is that the former is a fixed-sample property (if an estimator is unbiased, it is unbiased for any fixed sample size), while the latter is an asymptotic property (i.e., it is concerned only with what happens as the sample size becomes very large).

3.1.3 Efficiency

Between two statistics G and H both of which are assumed to be unbiased estimators of a parameter θ, G is the more "efficient estimator" if it has a smaller variance than G. This concept of efficiency is restricted to unbiased estimators only. A more general concept is that of "minimum mean-squared error". If G is an estimator of θ, the mean-square error of G is

$$MSE(G) = E(G-\theta)^2 . \tag{3.3}$$

Note that if G is unbiased, the mean-square error is identical to the variance of G.

3.1.4 Sufficiency

A statistic G is said to be a "sufficient" estimator of the parameter θ if G contains all of the relevant information available in the data about the value of θ. In other words a statistic satisfies the criterion of sufficiency when no other statistics which can be calculated from the same sample provides any additional information as to the value of the parameter to be estimated.

For example, if we have a set of observations x that we know comes from a normal population, sufficient statistics are Σx_i andΣx_i^2. No other information is needed to estimate μ and σ^2. Several theorems used to determine the sufficient statistic are discussed in details in Ref. [4].

3.2 Generation of Simulated Failure Data [19,20]

To construct the distribution of estimators by empirical means, it is first necessary to produce a large number of failure data pairs (i.e., number of failures k out of number of tries n) in which the number of failures k are distributed according to a known beta-binomial distribution with parameters a, b and n. The failure data pairs can then be grouped into sets of specified sizes and each set used with the various parameter estimation techniques to obtain parameter estimators. By analyzing a large number of data sets, the distribution of the estimators for each estimation method can then be determined empirically.

The sequence for generating synthetic failure data is performed by the following two steps:

1. A number of tries n is selected randomly from a discrete uniform distribution between n_1 and n_2.

2. Then, with the value of n from step 1, the number of failure k is chosen from a beta-binomial distribution

$$h(k;n,a,b) = \binom{n}{k} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+k)\Gamma(b+n-k)}{\Gamma(a+b+n)}, \qquad (3.4)$$

where a and b are specified parameters of the beta-binomial distribution. This procedure is simply a random sampling from a discrete uniform distribution and a beta-binomial distribution.

One widely-used concept of random sampling technique is the "inverse transformation". This method makes use of transformation of a "random number" from a uniform distribution between 0 and 1 inclusively to a random variable from a specified distribution. Once a random number is generated, the desired random variable can be obtained by the appropriate transformation. To find this transformation, suppose f(x) is a probability density function for a required random variable. The associated cumulative probability function F(x) is

$$F(x) = \int_{-\infty}^{x} f(s)\, ds. \qquad (3.5)$$

In particular, the cumulative probability function of a uniform distribution ($g(u) = 1$ for u between 0 and 1) is

$$G(u) = \int_0^u g(v)\, dv = u. \tag{3.6}$$

A random variable u which is described by a uniform distribution g(u) can be easily generated from any random number generating computer routine, such as RANDU [21], therefore a value of u or G(u) is known. A known value of u is transformed to a random variable x which is governed by a probability density function f(x) by using the fact that x and u have the same probability of being observed or a cumulative probability function F(x) is equal to G(u), i.e.,

$$\int_{-\infty}^{x} f(x')dx' = \int_0^u g(u')du', \tag{3.7}$$

or

$$\int_{-\infty}^{x} f(x')dx' = u. \tag{3.8}$$

In Eq. (3.8), f(x') and u are known. Therefore by solving Eq. (3.8) for x, the required random variable is obtained. The details of variable transformations from a uniform distribution to number of failures and number of tries will be presented in the next two sections.

3.2.1 Generation of Number of Tries, n [20]

In this study, n is assumed to be uniformly distributed between two positive integers M_1 and M_2. To obtain each value of n, a random number u, is generated from the routine RANDU [21] which is then transformed to value of n by the following algorithm:

$$n = \begin{cases} M_1 + \text{integer}[u/p] & , u \neq p \\ M_1 + \text{integer}[u/p]-1 & , u=p \end{cases} \quad (3.9)$$

where $p = (M_2-M_1-1)^{-1}$ (the probability of obtaining any integer between M_1 and M_2 inclusively). Explicitly, the above algorithm is equivalent to

$$n = \begin{cases} M_1 & , \quad 0 \leq u \leq p \\ M_{1+1} & , \quad p < u \leq 2p \\ \cdot & \quad\quad \cdot \\ \cdot & \quad\quad \cdot \\ M_{1+i} & , \quad ip < u \leq (i+1)p \\ \cdot & \quad\quad \cdot \\ \cdot & \quad\quad \cdot \\ \cdot & \quad\quad \cdot \\ M_2 & , \quad (1-p) < u \leq 1 \quad . \end{cases} \quad (3.10)$$

3.2.2 Generation of Number of Failures k

Once the number of tries, n, is selected randomly from a uniform distribution between M_1 and M_2, a new random number, u is generated and changed to the number of failures, k, by making use of the cumulative distribution, F(k), for a beta-binomial distribution, h(k), from Eq. (3.11), i.e.,

$$F(k) = \sum_{m=0}^{k} h(m;n,a,b), \quad k=0,1,\ldots,n. \quad (3.11)$$

A value of k is computed in such a way that k is the minimum integer for which $u \leq F(k)$, or equivalently,

$$k = \begin{cases} 0 \quad , & 0 \leq u \leq F(0) \\ 1 \quad , & F(0) < u \leq F(1) \\ \cdot & \cdot \\ \cdot & \cdot \\ i \quad , & F(i-1) < u \leq F(i) \\ \cdot & \cdot \\ \cdot & \cdot \\ n \quad , & F(u-1) < u \leq F(n) = 1 \end{cases} \tag{3.12}$$

In essence this method requires sequential evaluation of the cumulative distribution, F(k). Use of Eq. (3.11) for each calculation would be very time-consuming if large amounts of simulated failure data are to be generated. However, considerable computation effort may be saved in the sequential evaluation of F(k) by using the following recursion relations:

$$F(k+1) = F(k) + h(k+1;n,a,b) \ , \tag{3.13}$$

where

$$h(k+1;n,a,b) = h(k;n,a,b) \quad \frac{(a+k)(n-k)}{(b+n-k-1)(k+1)} \ . \tag{3.14}$$

In a low failure probability case for which a small value of k would be expected, the sequential search of Eq. (3.12) is best begun at k=0. Similarly, if a large value of k is expected (e.g., for a component with a high failure probability), then the search should be begun from k=n. More generally, to minimize the length of sequential search, the search should start near the mean of the beta-binomial distribution of interest. However, this optimal search method requires that the integer nearest to the mean, the density and cumulative functions [h(k;n,a,b) + F(k)] at that integer be initially evaluated and stored for all possible value of n. This refined search algorithm is outlined in Table 3.1.

Table 3.1. Algorithm for Optimal Calculation of Number of Failures, k, by the Inverse Transformation Technique.

```
Part I: Selection of Starting Values for Sequential Search

    1. Calculate means, $\mu_i$, of beta-binomials for all possible $n_i$
       (i.e., for $n_i = n_1, n_1+1, \ldots, n_2$).
    2. Round off means to nearest integer, $M_i$
    3. Calculate $F(M_i)$ and $h(M_i|n_i,a,b)$
    4. Store values of $M_i$, $F(M_i)$ and $h(M_i)$ in a vector to be used as
       starting points in sequential search.

Part II: Sequential Search to Calculate k for Given $n_i$

    1. Generate u from a uniform distribution on (0,1) by RANDU
    2. If $u = F(M_i)$, then $k = M_i$
    3. Otherwise, set $K = M_i$, $h(K) = h(M_i)$ and $F(K)=F(M_i)$
    4. If $u<F(M_i)$ go to step 6
    5. Compute:
           $h(K+1) = h(K) \frac{(a+K)(n_i-K)}{(b+n_i-K \quad (K+1)}$

           $F(K+1) = F(K) + h(K+1)$

       If $u \leq F(K+1)$ then $k = K+1$ and exit; otherwise set $K=K+1$
       and go back to beginning of step 5.

    6. Compute
           $F(K-1) = F(K) - h(K)$

       If $u>F(K-1)$ then $k=K-1$ and exit; otherwise calculate,
           $h(K-1) = h(K) \frac{K \, (n_i-K+b)}{(K-1+a)(n_i-K+1)}$

       set $K=K-1$, and go back to beginning of step 6.
```

4.0 RESULTS AND DISCUSSION

The main purpose of this study was to find from the distributions of estimators which of the five empirical estimation techniques described in Chapter 2 is the "best". To investigate the properties of the estimators from each method, simulated failure data were generated by methods described in the previous chapter and grouped into several different sample sizes. Based on these simulated data, estimates for the beta prior parameters from five methods were calculated by the five methods discussed in Chapter 3. With these estimates, their distributions were obtained empirically from which properties of estimators from each method were then investigated.

Since this study was concerned primarily with low failure probability events, a beta prior with parameters of $a=1.2$ and $b=23.0$ was used as the basis for generating the majority of the simulated failure data. (These particular values of a and b were the estimates calculated from observed data for emergency diesel generators.) The number of tries, n_i, was randomly selected from a discrete uniform distribution between 30 (M_1) and 300 (M_2) inclusively, using the technique described in Section 3.2.1. For each randomly chosen n_i, the associated number of failures, k_i, was selected randomly from a beta-binomial distribution with the parameters a and b specified above according to the procedure in Section 3.2.2. Many different values for n_i and k_i were produced in like manner. In all, 1,500 samples of size 5 (i.e., five pairs of k_i and n_i), 10 and 20 were generated. Additionally, 500 samples of size 50 were also computed. With these simulated data, estimates of the parameters a and b by the five estimation methods were calculated. As mentioned in Chapter 2, these methods did not always yield solutions for each sample. The observed number of samples yielding solution and percentage of success for each of the five methods are given in Table 4.1.

The PMM method which is the simplest estimation method always yielded successful results for all samples, regardless of sample size. All three of marginal-based methods occasionally failed, particularly for small sample sizes ($N \le 10$). However, for $N \geq 20$ all three marginal-based methods always yielded results. On the other hand, the PML method tended to fail increasingly as the sample size increased such that no parameter estimators were obtained for samples of size 50. All samples which failed to yield PML solutions contained at least one $k_i=0$ (a likely occurrence for the low failure probability case studied) which is the major deficiency of the prior maximum likelihood technique when used for low failure probability data.

Table 4.2 displays some simulation data samples for which no parameter estimators could be obtained by the marginal-based estimation techniques. No noticeable features about these particular data seem to distinguish them from other data samples for which the estimation methods yielded solutions.

4.1 Distribution of Prior Parameter Estimators

4.1.1 Mean and Variance

The frequency distributions of the estimates $\hat{a}$ and $\hat{b}$ as calculated by the various estimation techniques for the four sample sizes were con-

Table 4.1. Number of successful solutions and failures for prior parameter estimates from the simulation failure data for the five estimation techniques (outliers are considered as solutions).

Sample Size	WMMM Sol.	WMMM No-Sol.	WMMM % Success	Prior Matching Mom. Sol.	Prior Matching Mom. No-Sol.	Prior Matching Mom. % Success
5	1383	117	92.20	1500	0	100.0
10	1499	1	99.93	1500	0	100.0
20	1500	0	100.0	1500	0	100.0
50	500	0	100.0	500	0	100.0

Sample Size	Marg. Max. Likelihood Sol.	Marg. Max. Likelihood No-Sol.	Marg. Max. Likelihood % Success	Prior Max. Likelihood Sol.	Prior Max. Likelihood No-Sol.	Prior Max. Likelihood % Success
5	1349	151	89.93	850	650	56.67
10	1497	3	99.80	466	1034	31.07
20	1500	0	100.0	157	1343	10.47
50	500	0	100.0	0	500	0.00

Sample Size	UWMMM Sol.	UWMMM No-Sol.	UWMMM % success
5	1415	85	94.30
10	1497	3	99.80
20	1500	0	100.0
50	500	0	100.0

Table 4.2. Simulated failure data $\begin{pmatrix} n_i \\ k_i \end{pmatrix}$ from a beta-binomial ($a=1.2, b=23$) for which the marginal-based estimation methods yielded no solution.

Sample Size N=5

1. Data for which MML, WMMM, and UWMMM methods gave no solution:

$$\begin{pmatrix} 82 & 72 & 72 & 44 & 178 \\ 4 & 3 & 2 & 1 & 7 \end{pmatrix} \qquad \begin{pmatrix} 80 & 175 & 146 & 207 & 164 \\ 2 & 7 & 3 & 4 & 8 \end{pmatrix}$$

$$\begin{pmatrix} 235 & 156 & 46 & 112 & 90 \\ 8 & 4 & 1 & 3 & 0 \end{pmatrix} \qquad \begin{pmatrix} 95 & 53 & 125 & 299 & 69 \\ 1 & 0 & 3 & 2 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 285 & 247 & 152 & 237 & 228 \\ 5 & 2 & 2 & 2 & 6 \end{pmatrix} \qquad \begin{pmatrix} 72 & 186 & 243 & 105 & 73 \\ 1 & 7 & 4 & 5 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 232 & 72 & 82 & 238 & 105 \\ 20 & 3 & 6 & 24 & 10 \end{pmatrix} \qquad \begin{pmatrix} 271 & 191 & 197 & 270 & 283 \\ 5 & 2 & 3 & 3 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 209 & 207 & 76 & 62 & 117 \\ 10 & 8 & 2 & 4 & 2 \end{pmatrix} \qquad \begin{pmatrix} 254 & 225 & 298 & 120 & 220 \\ 10 & 10 & 19 & 6 & 12 \end{pmatrix}$$

2. Data for which only MML and WMMM methods failed:

$$\begin{pmatrix} 108 & 104 & 84 & 48 & 230 \\ 0 & 3 & 4 & 1 & 8 \end{pmatrix} \qquad \begin{pmatrix} 153 & 242 & 289 & 49 & 30 \\ 4 & 11 & 18 & 3 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 298 & 208 & 149 & 225 & 246 \\ 8 & 1 & 5 & 4 & 7 \end{pmatrix} \qquad \begin{pmatrix} 261 & 148 & 60 & 197 & 223 \\ 22 & 5 & 6 & 17 & 17 \end{pmatrix}$$

$$\begin{pmatrix} 91 & 72 & 165 & 161 & 87 \\ 1 & 4 & 5 & 2 & 2 \end{pmatrix} \qquad \begin{pmatrix} 212 & 117 & 120 & 39 & 151 \\ 6 & 3 & 7 & 0 & 6 \end{pmatrix}$$

$$\begin{pmatrix} 239 & 222 & 76 & 182 & 79 \\ 12 & 16 & 1 & 12 & 5 \end{pmatrix} \qquad \begin{pmatrix} 59 & 58 & 41 & 68 & 51 \\ 2 & 0 & 0 & 3 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 69 & 85 & 291 & 246 & 91 \\ 5 & 3 & 13 & 11 & 1 \end{pmatrix} \qquad \begin{pmatrix} 136 & 294 & 259 & 56 & 111 \\ 6 & 14 & 13 & 0 & 4 \end{pmatrix}$$

3. Data for which only WMMM and UWMMM methods failed:

$$\begin{pmatrix} 253 & 32 & 39 & 150 & 97 \\ 2 & 2 & 2 & 10 & 4 \end{pmatrix} \qquad \begin{pmatrix} 208 & 60 & 33 & 253 & 151 \\ 6 & 2 & 7 & 1 & 10 \end{pmatrix}$$

$$\begin{pmatrix} 213 & 89 & 209 & 248 & 122 \\ 9 & 3 & 5 & 3 & 2 \end{pmatrix} \qquad \begin{pmatrix} 130 & 53 & 247 & 155 & 45 \\ 6 & 1 & 3 & 4 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 85 & 87 & 123 & 269 & 63 \\ 3 & 6 & 7 & 6 & 3 \end{pmatrix} \qquad \begin{pmatrix} 166 & 59 & 61 & 104 & 150 \\ 5 & 4 & 3 & 6 & 12 \end{pmatrix}$$

$$\begin{pmatrix} 36 & 249 & 207 & 99 & 172 \\ 1 & 5 & 0 & 2 & 4 \end{pmatrix} \qquad \begin{pmatrix} 92 & 263 & 225 & 71 & 146 \\ 3 & 18 & 11 & 2 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 38 & 128 & 46 & 175 & 223 \\ 0 & 2 & 1 & 1 & 7 \end{pmatrix} \qquad \begin{pmatrix} 237 & 61 & 77 & 227 & 47 \\ 2 & 0 & 1 & 7 & 1 \end{pmatrix}$$

Table 4.2 - continued

4. Data for which only MML and UWMMM methods failed:

None

5. Data for which only the WMMM method failed:

$$\begin{pmatrix} 246 & 249 & 227 & 167 & 225 \\ 12 & 13 & 4 & 8 & 14 \end{pmatrix} \qquad \begin{pmatrix} 193 & 192 & 292 & 277 & 264 \\ 11 & 8 & 22 & 11 & 12 \end{pmatrix}$$

$$\begin{pmatrix} 119 & 150 & 133 & 282 & 262 \\ 5 & 7 & 2 & 12 & 5 \end{pmatrix} \qquad \begin{pmatrix} 86 & 216 & 139 & 189 & 33 \\ 0 & 4 & 0 & 3 & 0 \end{pmatrix}$$

6. Data for which only MML method failed:

$$\begin{pmatrix} 184 & 63 & 42 & 48 & 196 \\ 6 & 5 & 0 & 4 & 10 \end{pmatrix} \qquad \begin{pmatrix} 124 & 183 & 229 & 200 & 240 \\ 7 & 6 & 3 & 7 & 8 \end{pmatrix}$$

$$\begin{pmatrix} 169 & 44 & 63 & 260 & 182 \\ 15 & 2 & 4 & 0 & 15 \end{pmatrix} \qquad \begin{pmatrix} 239 & 281 & 168 & 183 & 191 \\ 20 & 18 & 14 & 6 & 14 \end{pmatrix}$$

$$\begin{pmatrix} 219 & 160 & 202 & 249 & 31 \\ 3 & 3 & 2 & 3 & 3 \end{pmatrix} \qquad \begin{pmatrix} 269 & 40 & 81 & 118 & 169 \\ 24 & 3 & 2 & 14 & 19 \end{pmatrix}$$

$$\begin{pmatrix} 298 & 234 & 115 & 209 & 113 \\ 10 & 9 & 2 & 13 & 3 \end{pmatrix} \qquad \begin{pmatrix} 214 & 82 & 116 & 84 & 225 \\ 0 & 5 & 11 & 4 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 35 & 72 & 45 & 130 & 50 \\ 1 & 2 & 2 & 4 & 5 \end{pmatrix} \qquad \begin{pmatrix} 67 & 38 & 133 & 133 & 268 \\ 0 & 3 & 3 & 6 & 6 \end{pmatrix}$$

7. Data for which only UWMMM method failed:

$$\begin{pmatrix} 156 & 58 & 159 & 276 & 292 \\ 1 & 1 & 2 & 10 & 5 \end{pmatrix} \qquad \begin{pmatrix} 60 & 256 & 40 & 37 & 261 \\ 3 & 4 & 3 & 3 & 22 \end{pmatrix}$$

$$\begin{pmatrix} 86 & 122 & 289 & 245 & 281 \\ 2 & 3 & 1 & 7 & 9 \end{pmatrix} \qquad \begin{pmatrix} 39 & 64 & 152 & 236 & 288 \\ 2 & 3 & 3 & 19 & 11 \end{pmatrix}$$

$$\begin{pmatrix} 49 & 284 & 67 & 136 & 271 \\ 2 & 5 & 2 & 4 & 17 \end{pmatrix} \qquad \begin{pmatrix} 173 & 144 & 201 & 94 & 48 \\ 11 & 2 & 11 & 4 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 39 & 219 & 203 & 276 & 202 \\ 1 & 2 & 6 & 9 & 10 \end{pmatrix} \qquad \begin{pmatrix} 133 & 63 & 116 & 252 & 287 \\ 6 & 2 & 2 & 7 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 127 & 167 & 53 & 147 & 227 \\ 6 & 5 & 2 & 11 & 7 \end{pmatrix} \qquad \begin{pmatrix} 56 & 36 & 137 & 63 & 34 \\ 5 & 3 & 4 & 4 & 4 \end{pmatrix}$$

Table 4.2 - continued

Sample Size N=10

1. Sample for which only WMMM and UWMMM methods found no solution:

$$\begin{pmatrix} 225 & 85 & 73 & 71 & 238 & 167 & 245 & 91 & 187 & 67 \\ 7 & 1 & 2 & 0 & 7 & 4 & 4 & 1 & 0 & 1 \end{pmatrix}$$

2. Sample for which only UWMMM method failed:

$$\begin{pmatrix} 63 & 195 & 34 & 293 & 42 & 70 & 276 & 295 & 264 & 74 \\ 3 & 1 & 1 & 15 & 2 & 3 & 10 & 10 & 2 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 170 & 221 & 241 & 236 & 178 & 45 & 244 & 201 & 69 & 123 \\ 5 & 13 & 8 & 7 & 10 & 1 & 5 & 5 & 3 & 6 \end{pmatrix}$$

3. Sample for which only MML method failed:

$$\begin{pmatrix} 40 & 111 & 108 & 273 & 217 & 207 & 254 & 31 & 284 & 108 \\ 0 & 6 & 8 & 14 & 8 & 10 & 7 & 0 & 14 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 36 & 206 & 254 & 97 & 95 & 99 & 276 & 233 & 253 & 281 \\ 1 & 10 & 8 & 3 & 7 & 10 & 15 & 11 & 11 & 12 \end{pmatrix}$$

$$\begin{pmatrix} 152 & 86 & 85 & 206 & 75 & 88 & 267 & 279 & 111 & 229 \\ 10 & 8 & 4 & 12 & 11 & 4 & 18 & 11 & 3 & 16 \end{pmatrix}$$

structed empirically and the results are shown in Figs. 4.1 through 4.4. These frequency distributions show some common characteristics. All distributions exhibit a slowly decaying tail at high estimate values. Bias and variance decrease as the sample size increases. For small sample sizes ($N \leq 10$), there were obtained an appreciable number of inordinately large estimates or "outliers" from the marginal-based estimation techniques. Presence of these outliers caused some difficulties in the computation of statistics from the distribution of estimates since they are often orders of magnitude greater than the true parameter values. If the outliers are included, values of statistics of the estimator distributions would be determined principally by the outlier values. For example, the distribution for N=5 for estimator $\hat{a}$ from the MML method (Fig. 4.1) yields a mean $\bar{a} = 7.16$ and a variance $\mathrm{var}(\hat{a}) = 2581$ if all data are used, while if the 14 outliers ($\hat{a} > 100a$) are suppressed, a mean $\bar{a} = 3.79$ and a variance $\mathrm{var}(\hat{a}) = 59.4$ results (the true value of a is 1.2). In this study, those outliers which are greater than one hundred times the true values were classified together with those samples which yielded no solution and hence were not used in the computation of statistics unless explicitly specified to the contrary.

In Tables 4.3 and 4.4, means, variances and covariances of parameter estimates without outliers is presented along with the values when the outliers are included. The number of outliers, if any, is shown in square brackets in Table 4.3. The means of the parameter estimator distributions are always greater than the true parameter values except for the mean calculated

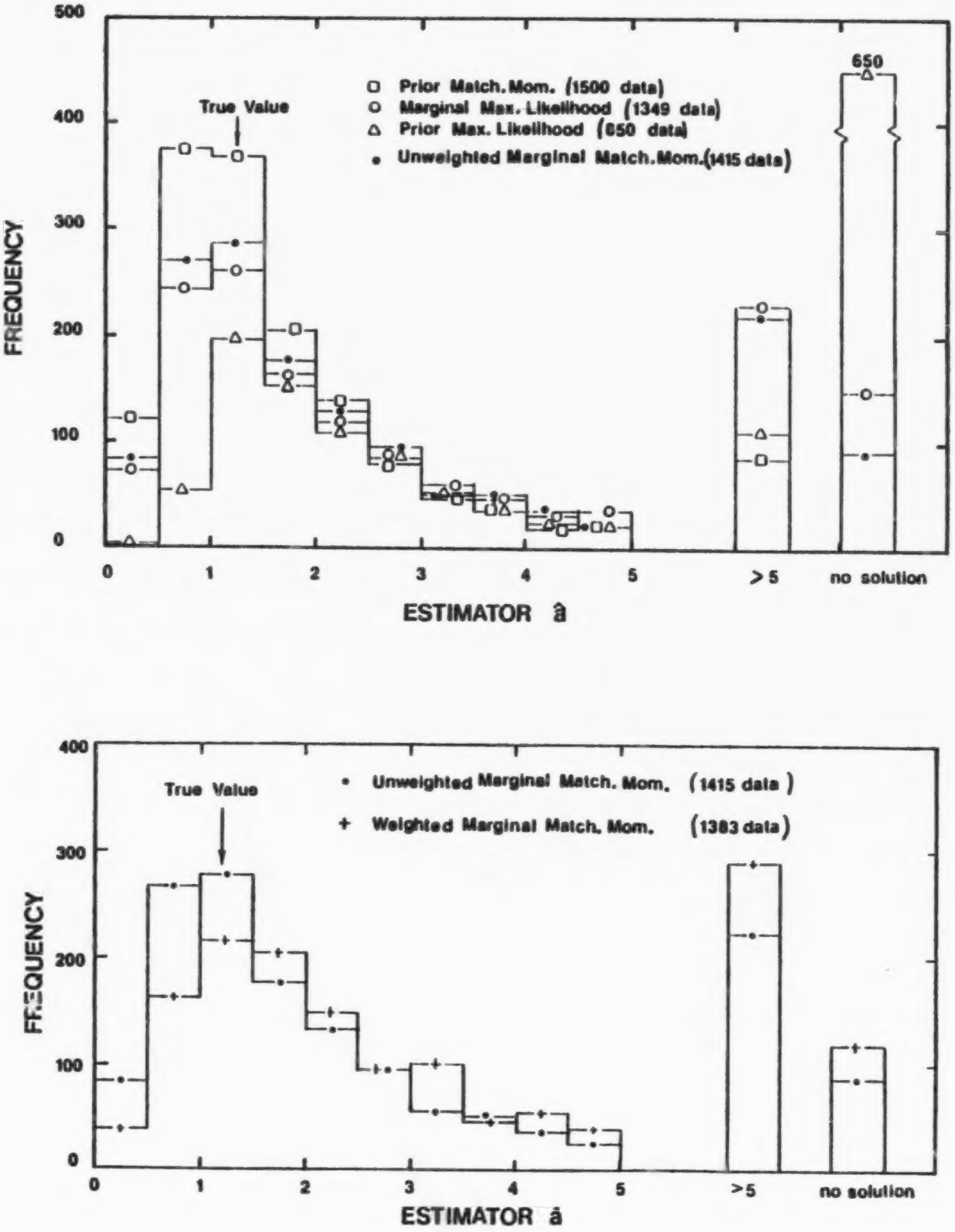


Fig. 4.1a Distribution of beta parameter estimators for samples of size N=5.

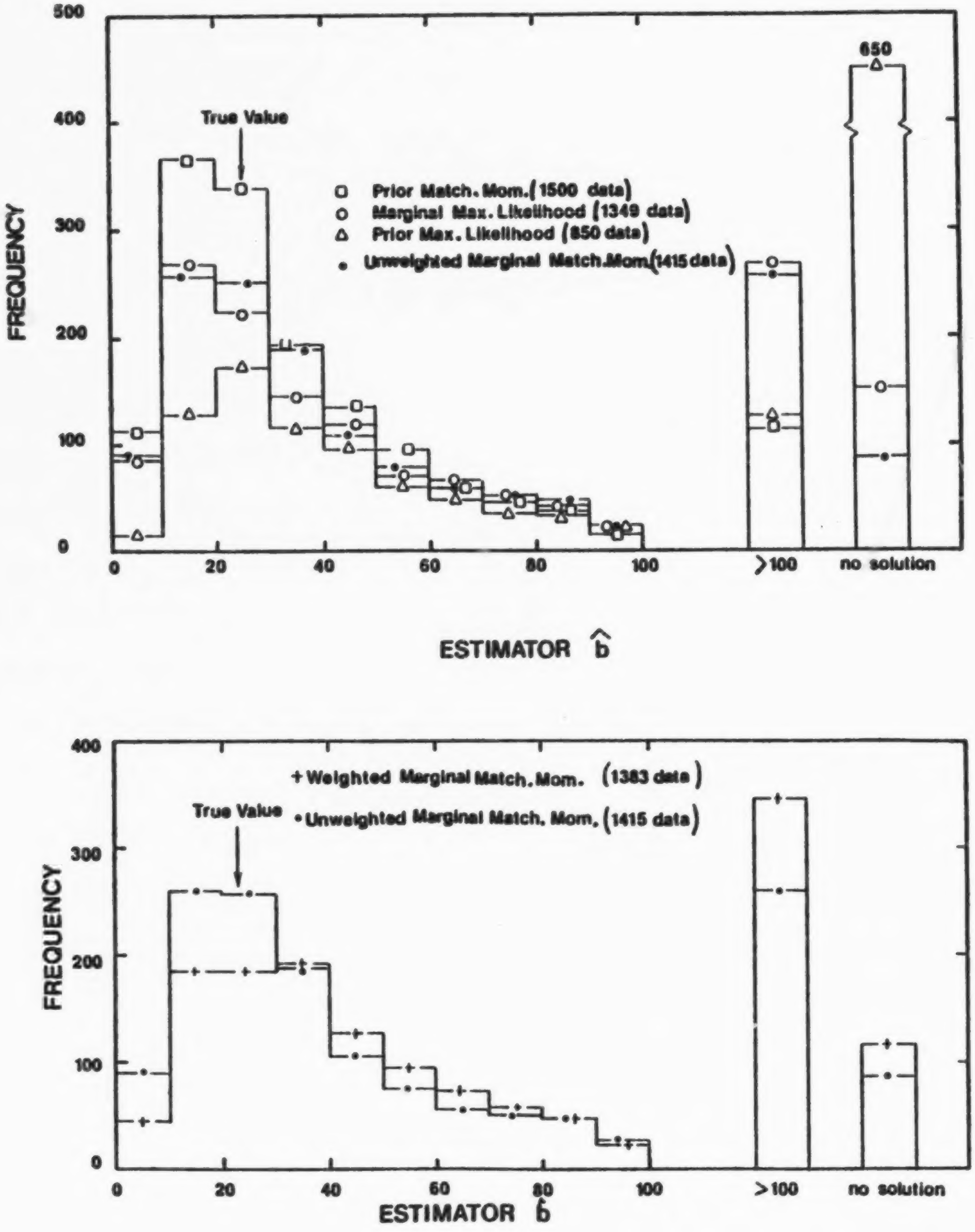


Fig. 4.1b Distribution of beta parameter estimators for samples of size N=5.

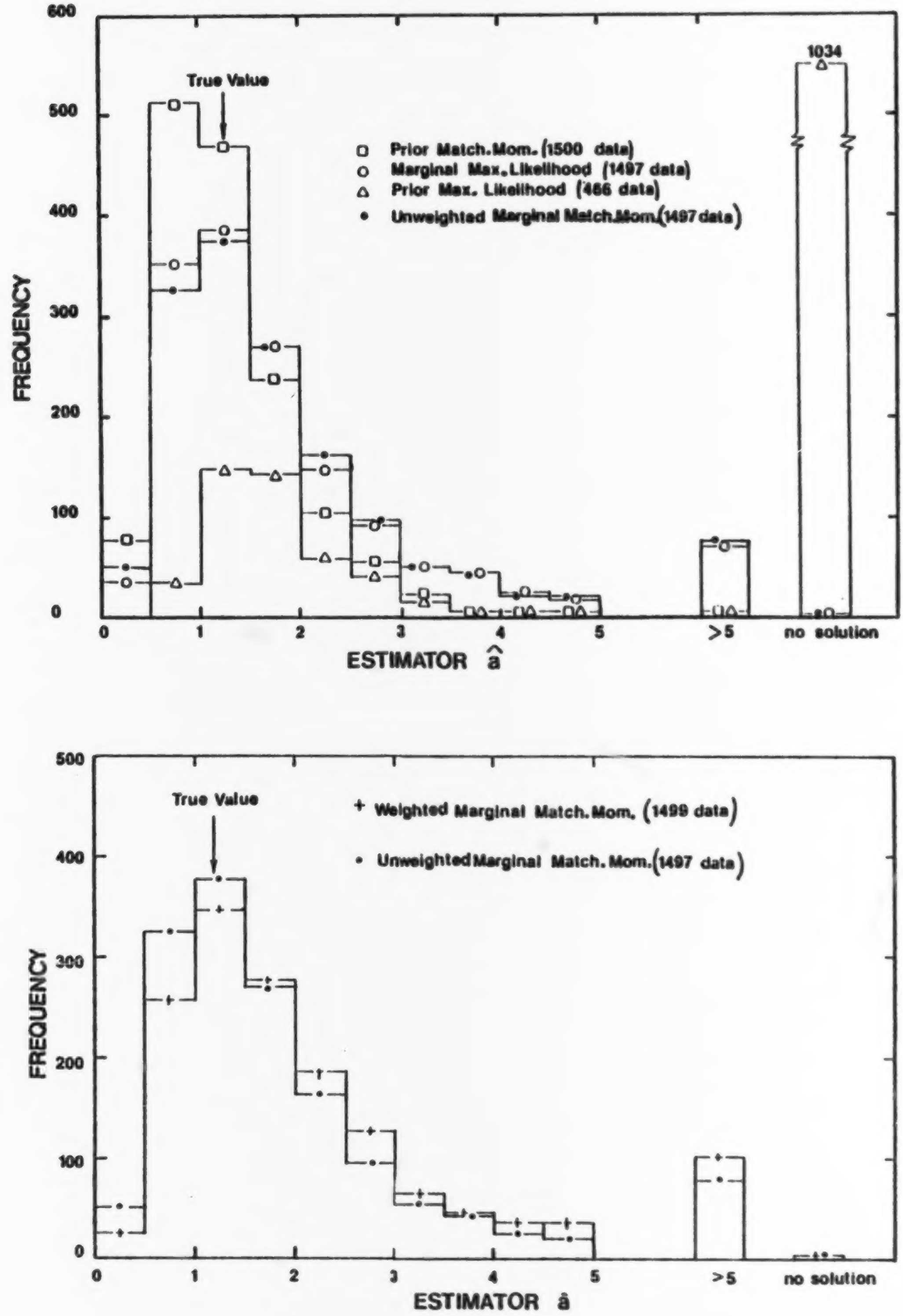


Fig. 4.2a Distribution of beta parameter estimators for samples of size N=10.

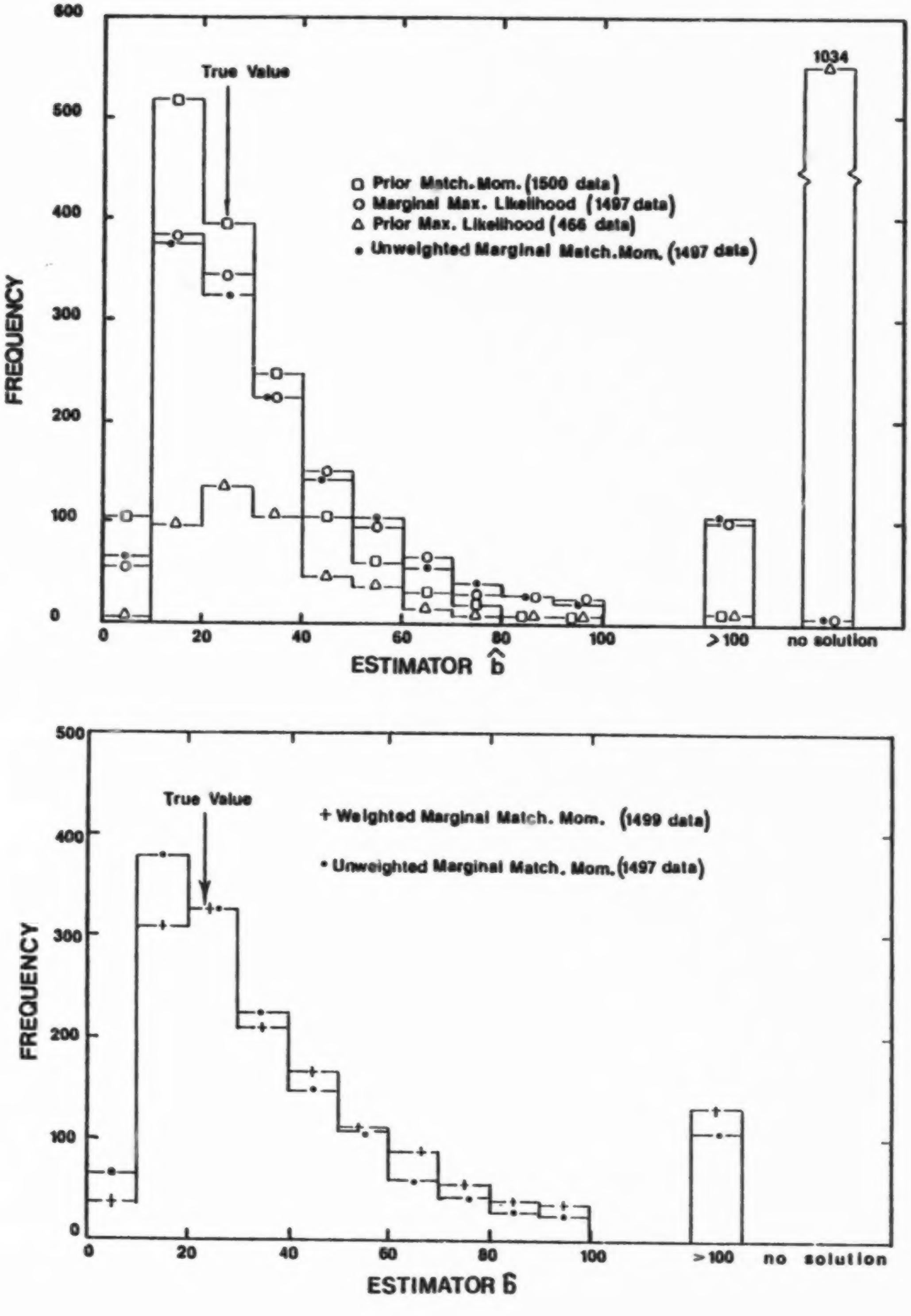


Fig. 4.2b Distribution of beta parameter estimators for samples of size N=10.

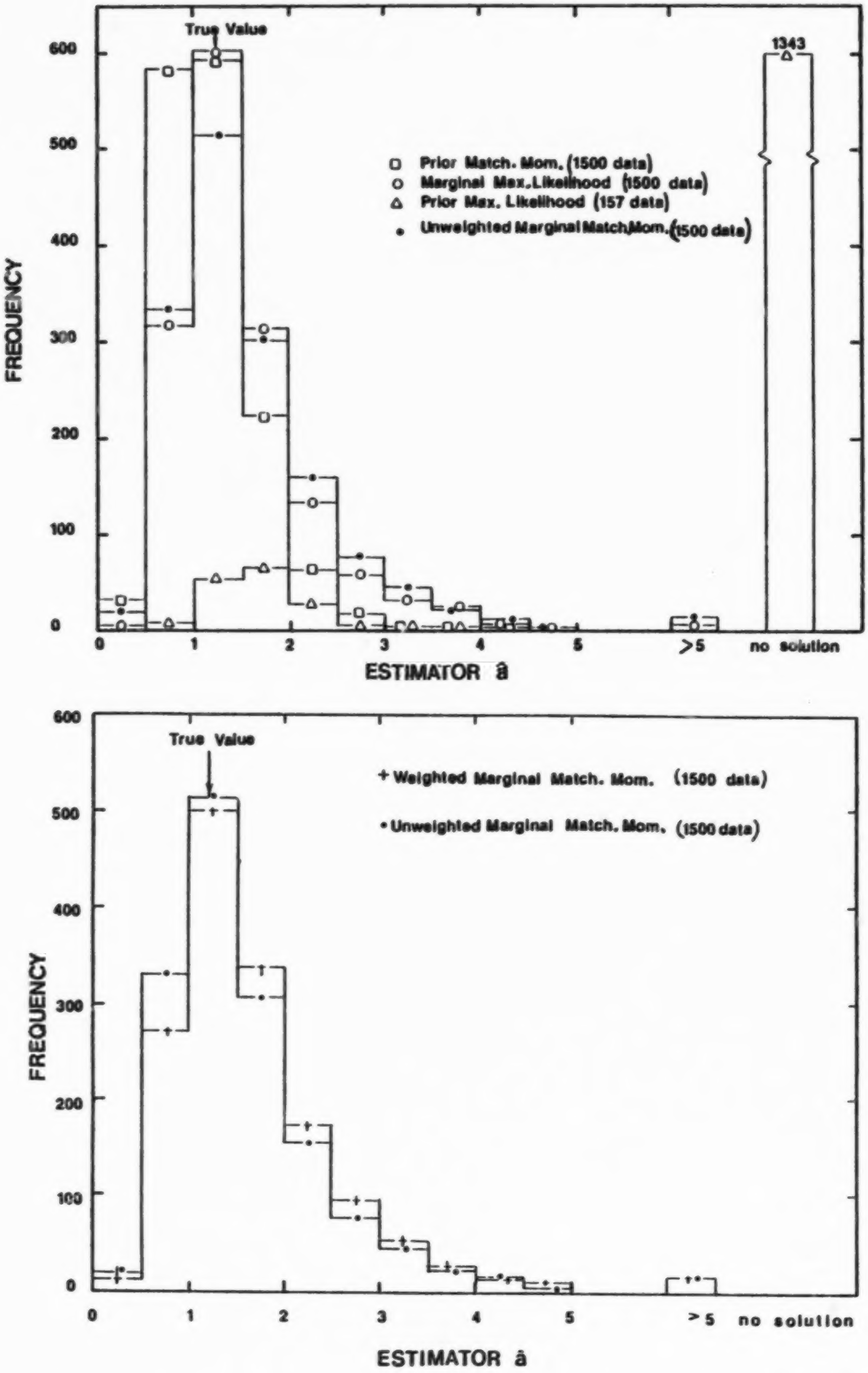


Fig. 4.3a Distribution of beta parameter estimators for samples of size N=20.

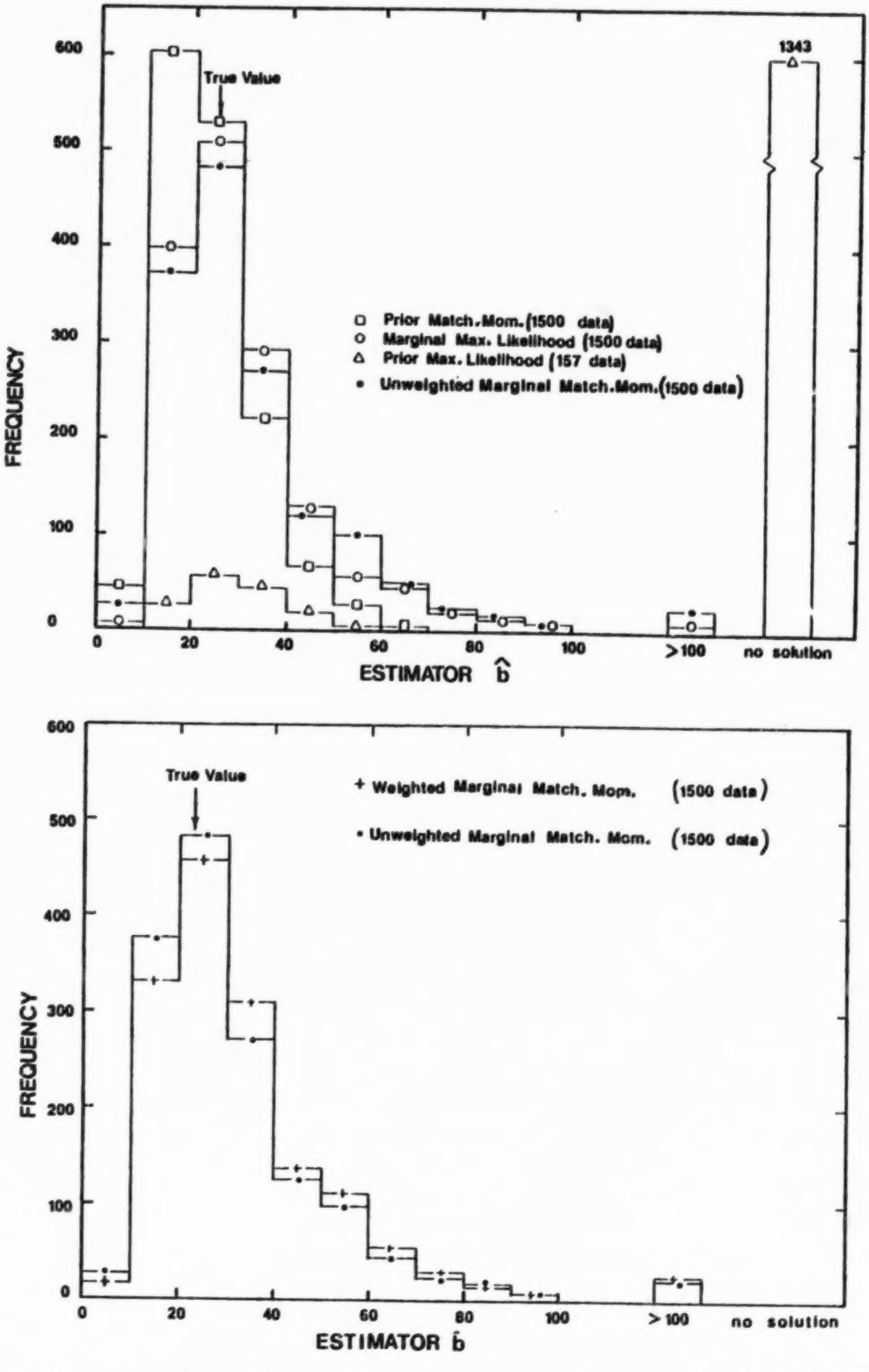


Fig. 4.3b Distribution of beta parameter estimators for samples of size N=20.

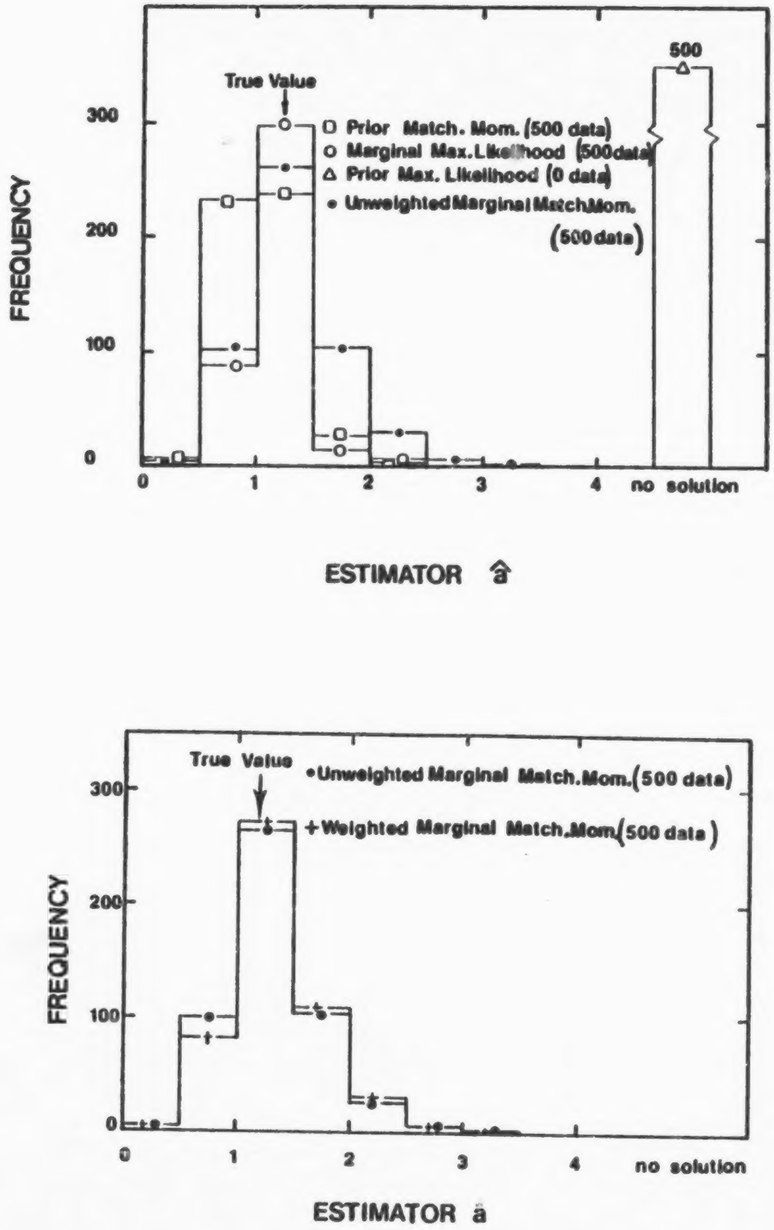


Fig. 4.4a Distribution of beta parameter estimators for samples of size N=50.

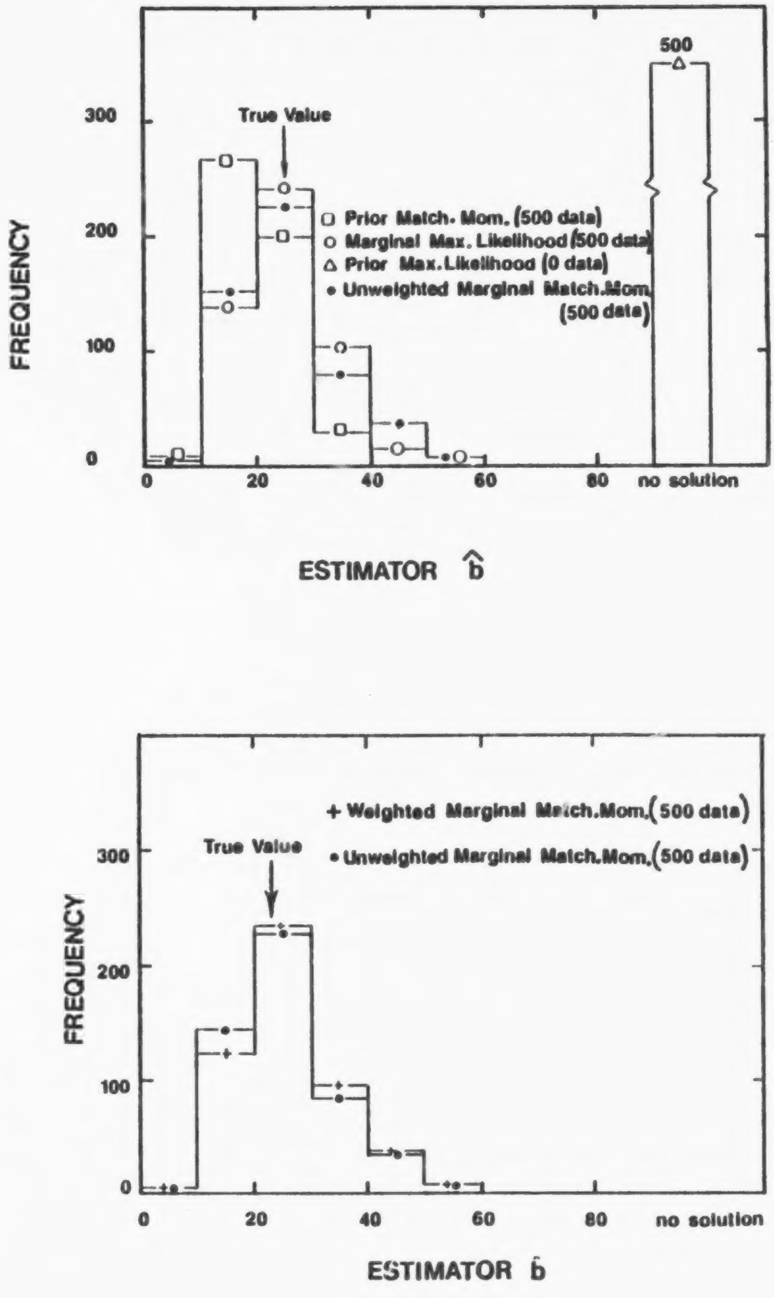


Fig. 4.4b Distribution of beta parameter estimators for samples of size N=50.

Table 4.3. Mean and number of solutions of parameter estimates for different sample sizes and estimation techniques. True beta parameter values are a=1.2 and b=23.0. Results for marginal-based methods are presented with and without outliers ($\hat{a} > 100$ a or $\hat{b}>100b$) included.

Sample Size	Prior Matching Moments			Prior Maximum Likelihood		
	$\hat{a}$	$\hat{b}$	# of sol.	$\hat{a}$	$\hat{b}$	# of sol.
5	1.91	43.9	1500	3.02	64.4	850
10	1.32	27.1	1500	1.87	34.8	466
20	1.18	23.4	1500	1.67	29.7	157
50	1.06	20.4	1500	-	-	0

Sample Size	WMMM w/o Outliers			WMMM with Outliers		
	$\hat{a}$	$\hat{b}$	# of sol.	$\hat{a}$	$\hat{b}$	# of sol.
5	4.28	93.3	1364	16.4	372.	1383 [19]*
10	2.43	51.7	1498	2.53	53.7	1499 [1]
20	1.68	34.0	1500	1.68	34.0	1500 [0]
50	1.36	26.4	500	1.36	26.4	500 [0]

Sample Size	Marg. Max. Like. w/o Outliers			Marg. Max. Like. with Outliers		
	$\hat{a}$	$\hat{b}$	# of sol.	$\hat{a}$	$\hat{b}$	# of sol.
5	3.79	83.7	1335	7.16	199.	1349 [14]*
10	2.04	44.3	1495	3.43	71.1	1497 [2]
20	1.52	30.8	1500	1.52	30.8	1500 [0]
50	1.30	25.2	500	1.30	25.2	500 [0]

Sample Size	WMMM w/o Outliers			UWMMM with outliers		
	$\hat{a}$	$\hat{b}$	# of sol.	$\hat{a}$	$\hat{b}$	# of sol.
5	3.55	80.4	1407	5.37	116.	1415 [8]*
10	2.20	47.0	1497	2.20	47.0	1497 [0]
20	1.60	32.2	1500	1.60	32.2	1500 [0]
50	1.33	25.8	500	1.33	25.8	500 [0]

*Figure in the bracket is number of outliers obtained from that sample size.

Table 4.4 Variances and covariance of parameter estimators for different sample sizes and estimation techniques. True beta parameter values are a=1.2 and b=23.0. Results for marginal-based methods are presented with and without outliers ($\hat{a}$>100a or $\hat{b}$>100b) included.

Sample Size	Prior Matching Moments			Prior Maximum Likelihood		
	var ($\hat{a}$)	var ($\hat{b}$)	cov ($\hat{a}$,$\hat{b}$)	var ($\hat{a}$)	var ($\hat{b}$)	cov ($\hat{a}$,$\hat{b}$)
5	4.42	3.79 (3)	1.03 (2)	9.29	8.44 (3)	2.23 (2)
10	5.50 (-1)*	2.86 (2)	1.01 (1)	8.40 (-1)	4.39 (2)	1.63 (1)
20	2.11 (-1)	9.97 (1)	3.79	2.02 (-1)	1.08 (2)	3.81
50	6.72 (-2)	3.05 (1)	1.23	-	-	-

Sample Size	WMMM w/o Outliers			WMMM with Outliers		
	var ($\hat{a}$)	var ($\hat{b}$)	cov ($\hat{a}$,$\hat{b}$)	var ($\hat{a}$)	var ($\hat{b}$)	cov ($\hat{a}$,$\hat{b}$)
5	5.20 (1)	2.50 (4)	9.90 (2)	8.15 (4)	3.41 (7)	1.64 (6)
10	1.23 (1)	5.76 (3)	2.51 (2)	2.69 (1)	1.15 (4)	5.40 (2)
20	8.01 (-1)	4.49 (2)	1.69 (1)	8.01 (-1)	4.49 (2)	1.69 (1)
50	1.75 (-1)	8.13 (1)	3.44	1.75 (-1)	8.13 (1)	3.44

Sample Size	Marg. Max. Like. w/o Outliers			Marg. Max. Like. with Outliers		
	var ($\hat{a}$)	var ($\hat{b}$)	cov ($\hat{a}$,$\hat{b}$)	var ($\hat{a}$)	var ($\hat{b}$)	cov ($\hat{a}$,$\hat{b}$)
5	5.94 (1)	2.74 (4)	1.15 (3)	2.58 (3)	6.39 (6)	1.18 (5)
10	5.60	4.09 (3)	1.37 (2)	2.89 (3)	1.08 (6)	5.59 (4)
20	5.70 (-1)	3.39 (2)	1.22 (1)	5.70 (-1)	3.39 (2)	1.22 (1)
50	1.14 (-1)	5.83 (1)	2.27	1.14 (-1)	5.83 (1)	2.27

Sample Size	UWMMM w/o Outliers			UWMMM with Outliers		
	var ($\hat{a}$)	var ($\hat{b}$)	cov ($\hat{a}$,$\hat{b}$)	var ($\hat{a}$)	var ($\hat{b}$)	cov ($\hat{a}$,$\hat{b}$)
5	5.00 (1)	2.78 (4)	1.07 (3)	1.44 (3)	3.30 (5)	1.97 (4)
10	1.35 (1)	6.58 (3)	2.82 (2)	1.35 (1)	6.58 (3)	2.82 (2)
20	8.54 (-1)	4.31 (2)	1.73 (1)	8.54 (-1)	4.31 (2)	1.73 (1)
50	1.71 (-1)	7.92 (1)	3.36	1.71 (-1)	7.92 (1)	3.36

*Read 5.50×10^{-1}.

from the results of the PMM method. For sample size 20, the mean from the PMM method is almost exactly the same as the true parameters, but the PMM method tends to underestimate for larger sample sizes.

Another important feature of these estimator distributions is their variances,i.e.,a measure dispersion of the estimators. For a given sample size the minimum variance was always obtained with the PMM method, which is also the simplest technique of those studied. These estimation methods based on the marginal distribution always yielded considerably larger variances, a result of the more slowly decaying tail of the distributions at large $\hat{a}$ and $\hat{b}$ values. Moreover, the covariances of $\hat{a}$ and $\hat{b}$ were always observed to be positive which indicates a positive linear relationship between $\hat{a}$ and $\hat{b}$ (i.e., large values of $\hat{a}$ are associated with larger values of $\hat{b}$). Finally, as would be expected, the variances and covariances for all estimation techniques decrease as the sample size increases, and the means approach the true values.

4.1.2 Bias of Prior Parameter Estimates

The degree of bias inherent in any parameter estimation technique is often of concern. The bias of an estimator, $\hat{\theta}$, is defined as,

$$\text{Bias} \equiv E[\hat{\theta}-\theta] = \bar{\theta}-\theta \qquad (4.2)$$

where θ is the true value of the parameter (e.g., a or b) and $\bar{\theta}$ is the mean of the estimators. All of the estimation techniques investigated in this study were found to yield biased estimates of the prior parameters, especially for small sample sizes.

In Table 4.5 the results are presented of the bias of the beta parameter estimators for each estimation method considered. The variation of

Table 4.5 The bias or deviation of mean of estimators from true parameters [a=1.2, b=23.0]. Each data set consists of 500 simulation samples.

Sample Size (N)	Data Set No.	PMM		MML		PML	
		$\bar{a}-a$	$\bar{b}-b$	$\bar{a}-a$	$\bar{b}-b$	$\bar{a}-a$	$\bar{b}-b$
5	1	0.566	16.5	2.76	63.1	1.49	2.0
	2	0.739	21.9	2.08	30.6	2.05	47.01
	3	0.835	2.41	2.91	68.4	1.95	45.6
10	1	0.124	3.72	0.887	21.2	0.673	10.6
	2	0.104	3.72	0.772	19.2	0.691	11.8
	3	0.125	4.82	0.872	23.5	0.660	13.0
20	1	-0.0238	0.0602	0.325	7.37	0.479	6.07
	2	-0.0574	0.299	0.268	6.71	0.439	6.48
	3	0.0118	1.44	0.373	9.34	0.491	7.38
50	1	-0.142	-2.58	0.100	2.22	*	*

Sample Size (N)	Data Set No.	WMMM		UWMMM	
		$\bar{a}-a$	$\bar{b}-b$	$\bar{a}-a$	$\bar{b}-b$
5	1	3.24	76.2	2.67	63.4
	2	2.68	61.8	1.77	42.2
	3	3.30	72.8	2.59	66.1
10	1	1.20	26.3	0.989	22.7
	2	1.12	26.5	0.966	23.6
	3	1.38	33.1	1.03	25.4
20	1	0.471	10.2	0.380	8.34
	2	0.412	9.50	0.318	7.63
	3	0.568	13.4	0.496	11.8
50	1	0.164	3.40	0.132	2.82

*Method always failed for sample size N=50 since each sample contained at least one k_i=0.

the bias in $\hat{a}$ and $\hat{b}$ with sample size is shown in Fig. 4.5. Notice that as the sample size increases, the bias of the estimator from the marginal maximum likelihood method decreases towards zero as would be expected from the consistency property of the maximum likelihood method. The estimates from both of the marginal matching moment techniques also exhibited this consistency property. The bias for the prior-based maximum likelihood method, on the other hand, is poor for large sample sizes since for the assumed prior beta many of the simulated samples contain at least one $k_i=0$ which makes this estimation method fail (see Table 4.1). However, it will be shown for the symmetric case (a=b=5.0) that the estimator from the prior maximum likelihood method can possess a consistency property when all samples are used.

From Fig. 4.5 it is seen that all of the methods except the simplest method, the PMM, always yield a positive bias. The bias of the prior matching moment results has a positive value for small sample sizes but becomes negative for sample size of greater than 20. More importantly, the PMM method has the smallest bias of all five methods investigated in this study for sample size (N) less than or equal to 50.

4.1.3 Mean Squared Error of Estimators

For safety analyses the mean squared error of an estimator is generally of concern. Although a particular method may have a small bias, the variance of the estimates may be quite large and hence the analysis of an individual sample could lead to parameter estimates which are significantly different from the true values. For safety considerations in which only a few samples are to be analyzed it is

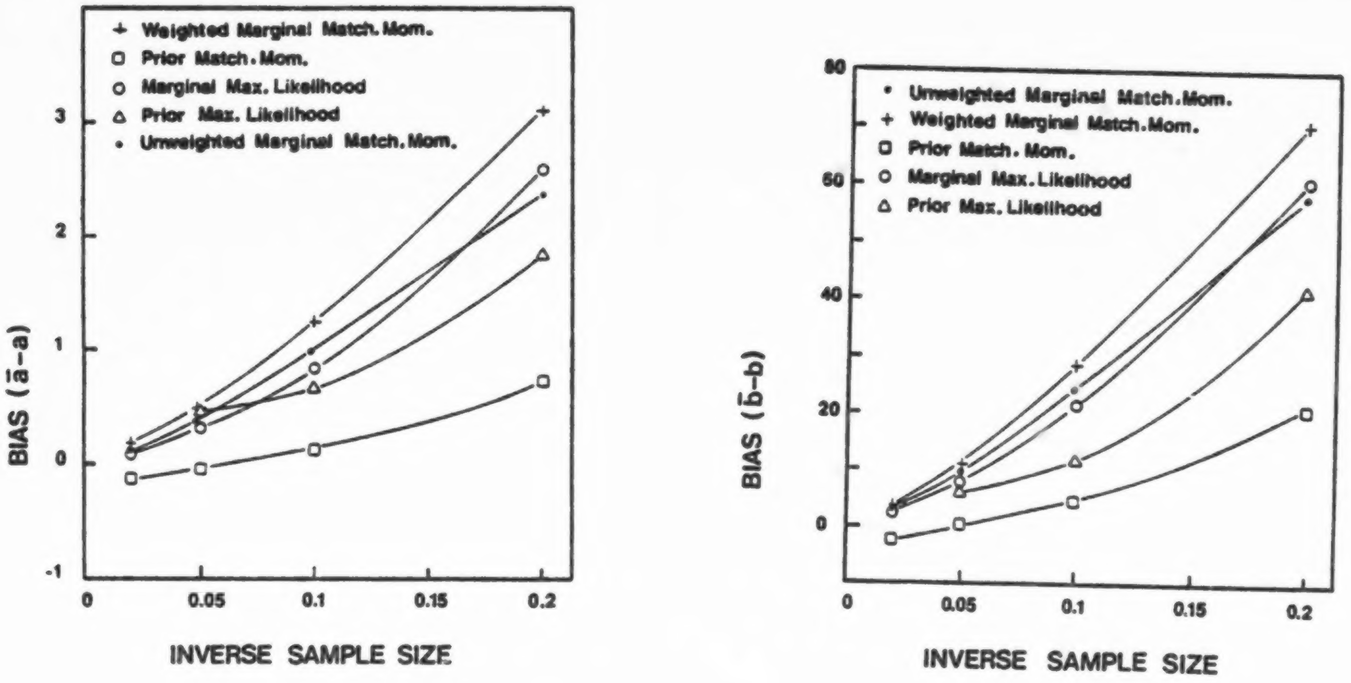


Fig. 4.5 Variation of the bias of the beta parameter estimators with sample size for the different estimation techniques. True values of the beta parameters are a=1.2 and b=23.

important that the mean square error of the estimates be small even if the estimates are slightly biased.

For the simulated data the mean squared error (MSE) is estimated as

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (\hat{\theta}_i - \theta)^2 \qquad (4.3)$$

where $\hat{\theta}_i$ represents the estimate $\hat{a}$ or $\hat{b}$ and θ represents the true value. From this equation, it is seen that outliers (i.e., estimates which are far removed from the true value) will change the value of the mean squared error greatly, and that estimates close to the true value have little influence. From the distributions of $\hat{a}$ and $\hat{b}$ shown in Figs. 4.1-4.4, it is seen that there are typically several outliers produced by the marginal-based estimation methods, especially for small sample sizes. To compare the mean squared error for the different estimation methods, these outliers were suppressed by ignoring those values of $\hat{a}$ or $\hat{b}$ which were more than one hundred times the true values of a and b. The results of the mean squared error analysis for the simulated failure data are presented in Table 4.6 and in Fig. 4.6.

From these results it is seen that for small or moderate sample sizes ($N \leq 50$) the prior matching moment estimation techniques yields the lowest mean squared error. The three estimation methods based on the marginal distribution produce the poorest results, i.e., the largest mean squared errors. These large errors are a direct result of the occasional high estimates of a and b obtained with these methods.

4.1.4 Median of Estimators

To suppress naturally the effect of outliers without actually

Table 4.6. Mean squared error about the true beta parameters (a=1.2, b=23) for the simulated failure data. Each data set contained 500 samples.

Sample Size (N)	Data Set No.	PMM		MML		PML	
		MSE($\hat{a}$)	MSE($\hat{b}$)	MSE($\hat{a}$)	MSE($\hat{b}$)	MSE($\hat{a}$)	MSE($\hat{b}$)
5	1	2.57	1,740	77.6	35,000	6.76	3,780
	2	6.43	4,860	43.4	21,000	17.1	11,500
	3	5.80	6,050	77.1	37,100	14.9	15,100
10	1	0.629	308	7.12	4,670	1.52	524
	2	0.526	290	4.68	3,090	1.20	566
	3	0.535	310	7.12	5,860	1.14	639
20	1	0.215	95.3	0.618	314	0.422	125
	2	0.185	89.1	0.680	382	0.288	128
	3	0.235	115	0.781	503	0.520	193
50	1	0.0874	37.1	0.123	63.1	-	-

Sample Size (N)	Data Set No.	WMMM		UWMMM	
		MSE($\hat{a}$)	MSE($\hat{b}$)	MSE($\hat{a}$)	MSE($\hat{b}$)
5	1	55.9	33,200	75.9	37,800
	2	58.0	27,500	39.7	19,200
	3	70.8	28,900	50.2	36,000
10	1	11.7	4,880	12.2	6,600
	2	7.61	4,480	14.6	7,610
	3	22.0	10,400	16.6	7,230
20	1	0.971	472	0.835	409
	2	0.806	455	0.688	384
	3	1.33	782	1.51	756
50	1	0.201	92.7	0.188	87.0

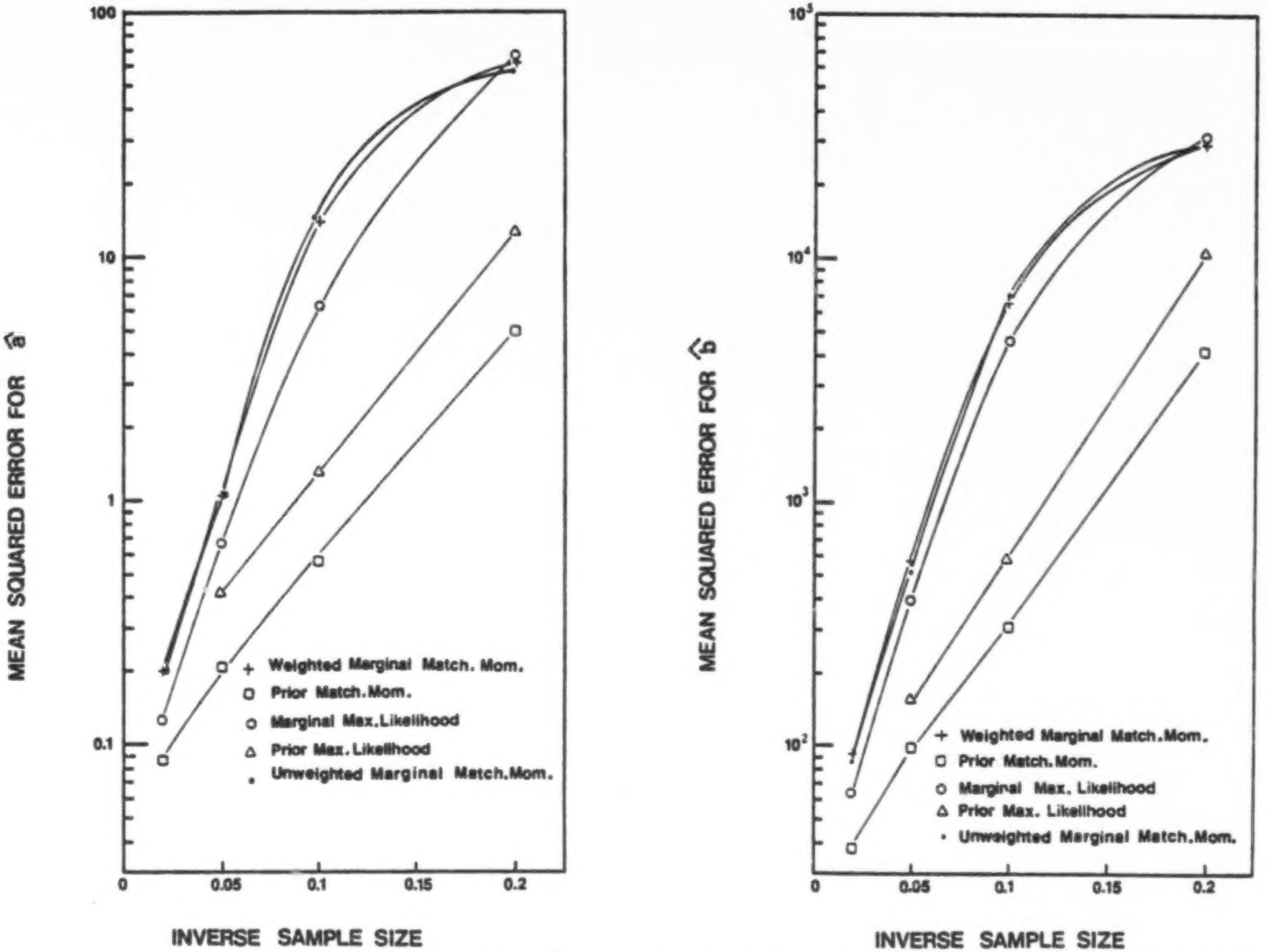


Fig. 4.6 Variation of the mean squared error of the beta parameter estimators with sample size for the different estimation techniques. True values of the beta parameters are a=1.2 and b=23.

ignoring them, the median of the empirical distributions for $\hat{a}$ and $\hat{b}$ were calculated. The results for the median of the distributions are given in Table 4.7 and the variation of the median with sample size is shown in Fig. 4.7. In the calculation of the median values, the outlier estimators were included. For small sample sizes ($N \leq 10$) the simple prior matching moments method yields median values which are closest to the true values of the parameters. However, for larger sample sizes the prior matching moment methods gives a median which is smaller than the true value. Only the estimation methods based on the marginal distribution appear to yield medians which approach the true value as the sample size becomes very large.

4.1.5 Comparison to Results from a Symmetric Beta Prior

The results in the previous section were estimated from simulation failure data based on a specific beta prior distribution which was highly skewed towards low failure probabilities (the mean of the beta prior $= a/(a+b) = 1.2/(1/2+23) = 0.043$). To determine whether or not the results obtained for the estimators of this particular beta prior are applicable only to similarly skewed beta priors or to more generally distributed beta priors, failure data were simulated for a symmetrically distributed beta prior with parameters $a=b=5$ and consequently with a mean of 0.5. Simulated failure data sets of 500 samples of size 5, 10 and 20 were generated from this symmetric beta distribution. The five estimation techniques were used to analyze these data.

From this analysis of failure data generated from a symmetric beta prior, it was found that three marginal-based estimation techniques

Table 4.7. Median values for the estimates $\hat{a}$ and $\hat{b}$ for different sample sizes and estimation techniques. For sample sizes of 5, 10 and 20, 1500 simulated failure data were used, and for sample size 50, 500 simulated data were used. The true value of the parameters are a=1.2 and b=23.0.

Sample Size (N)	WMMM		Prior Match. Mom.	
	$\hat{a}$	$\hat{b}$	$\hat{a}$	$\hat{b}$
5	2.22	46.3	1.31	27.8
10	1.72	33.5	1.76	23.0
20	1.47	28.4	1.10	21.4
50	1.28	24.4	1.02	19.6

Sample Size (N)	Marg. Max. Like.		Prior Max. Like.	
	$\hat{a}$	$\hat{b}$	$\hat{a}$	$\hat{b}$
5	1.77	36.9	2.09	39.2
10	1.47	28.9	1.65	29.9
20	1.33	25.6	1.67	29.2
50	1.23	23.3	-	-

Sample Size (N)	UWMMM	
	$\hat{a}$	$\hat{b}$
5	1.69	35.0
10	1.49	29.2
20	1.37	26.8
50	1.25	23.9

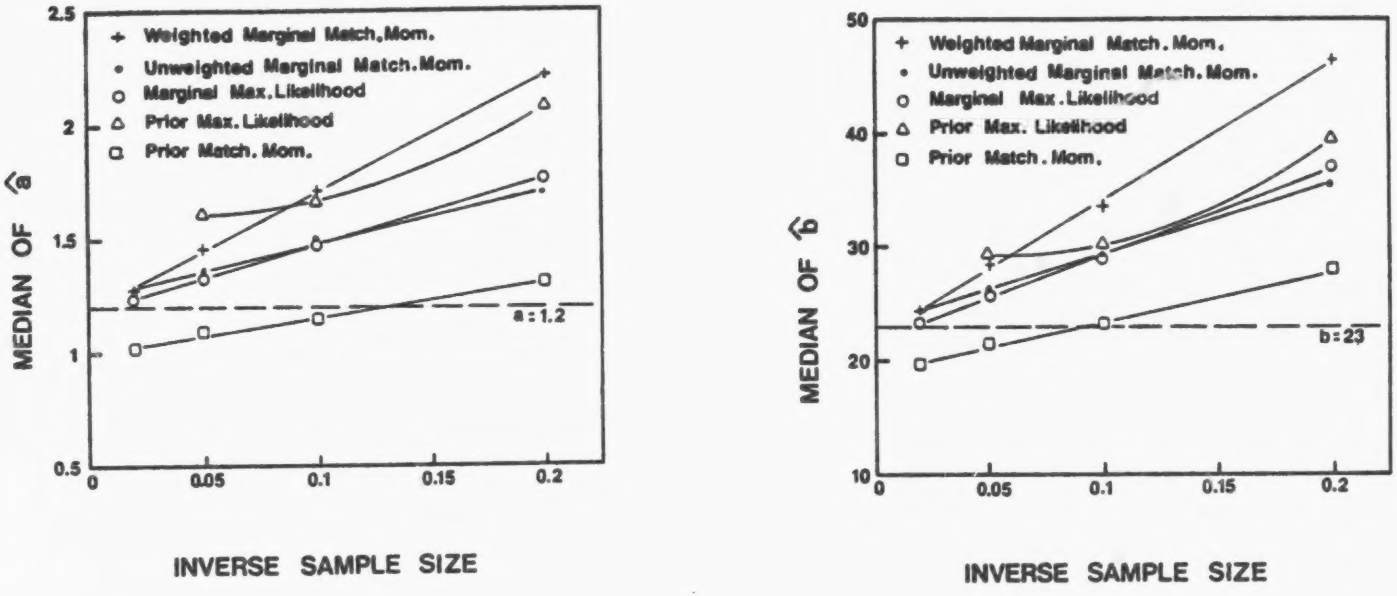


Fig. 4.7 Variation of the median of the beta parameter estimators with sample size for the different estimation techniques. True values of the beta parameters are a=1.2 and b=23.

yielded numerical solutions for a larger fraction of the samples than they did for the nonsymmetric case. For example, 98.8% of the size 5 samples yielded results with the weighted marginal matching method, 98.0% of the same samples were successfully analyzed by the marginal maximum likelihood method and 99.0% of success were obtained from the unweighted marginal matching moments techqiue. For the nonsymmetric case these success rate percentages were (see Table 4.1) 92.2%, 89.9% and 94.3%, respectively. Unlike the nonsymmetric case, no data samples of size greater than 5, which did not yield solutions, were found. Moreover, only two outliers were obtained, one from the weighted marginal matching moments ($\hat{a}$ = 636, $\hat{b}$ = 219) and the other from the unweighted marginal matching moments ($\hat{a}$ = 502, $\hat{b}$ = 396). In case of a symmetric beta prior, none of the simulated failure samples contained a k_i=0 (or k_i=n_i), and hence, unlike the skewed beta prior case, the prior maximum likelihood estimation method produced parameter estimates for all samples.

The results for the bias and the mean squared error of the estimetor produced parameter estimates for all samples. Figures 4.8 and 4.9 show the variation with sample size of the bias and mean squared error, respectively. Because the true beta parameters are equal (a=b=5), one would expect the plots of the bias for $\hat{a}$ to be the same as for $\hat{b}$. Indeed the small observed differences in Fig. 4.8 or in Table 4.8 are a result of statistical uncertainties arising from the relatively small number of samples (500) used to construct the distributions of $\hat{a}$ and $\hat{b}$.

From Fig. 4.8 all five methods appear to give zero or very small bias if the sample size becomes sufficiently large. As with the skewed

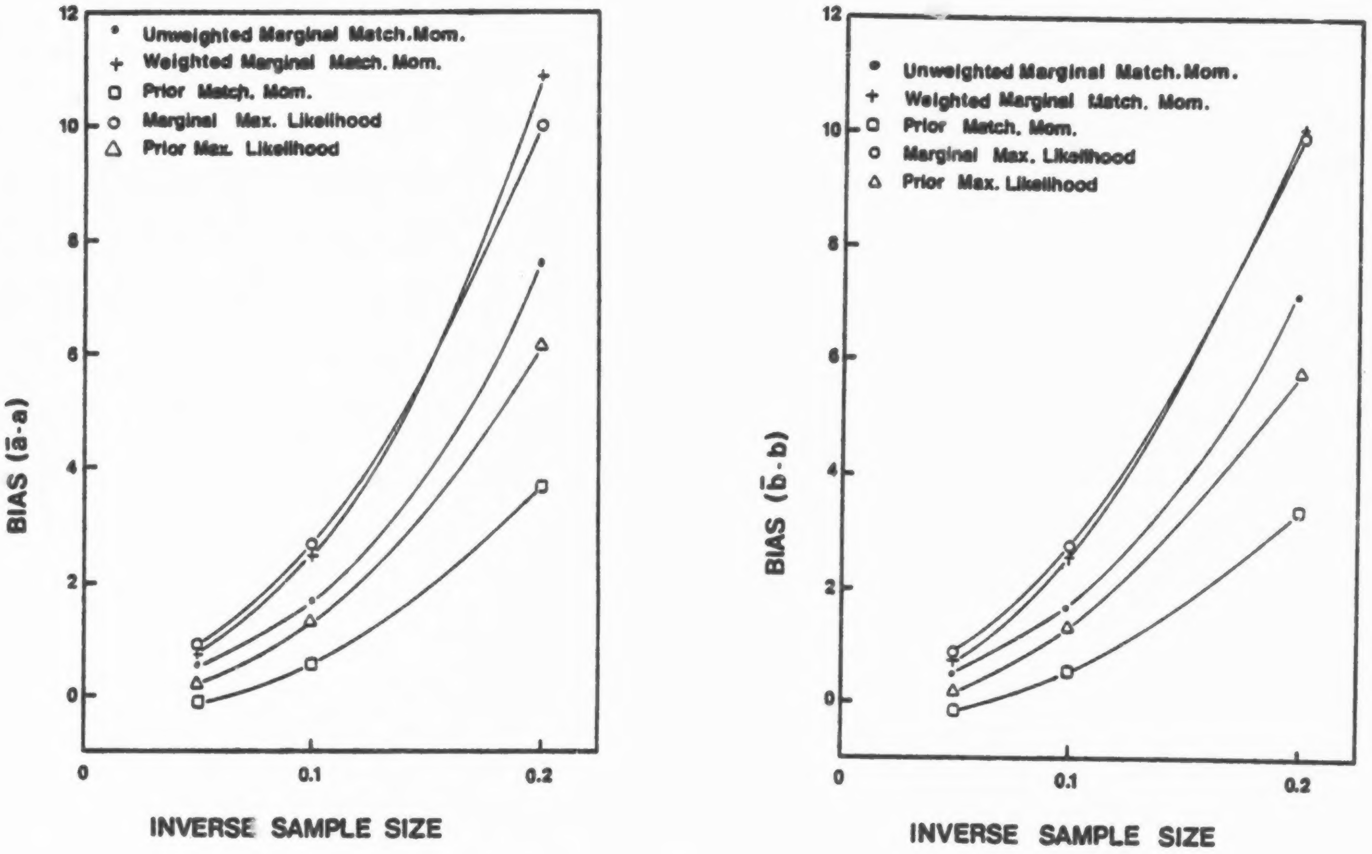


Fig. 4.8 Variation of the bias of the beta parameter estimators with sample size for the symmetric beta distribution (a=b=5).

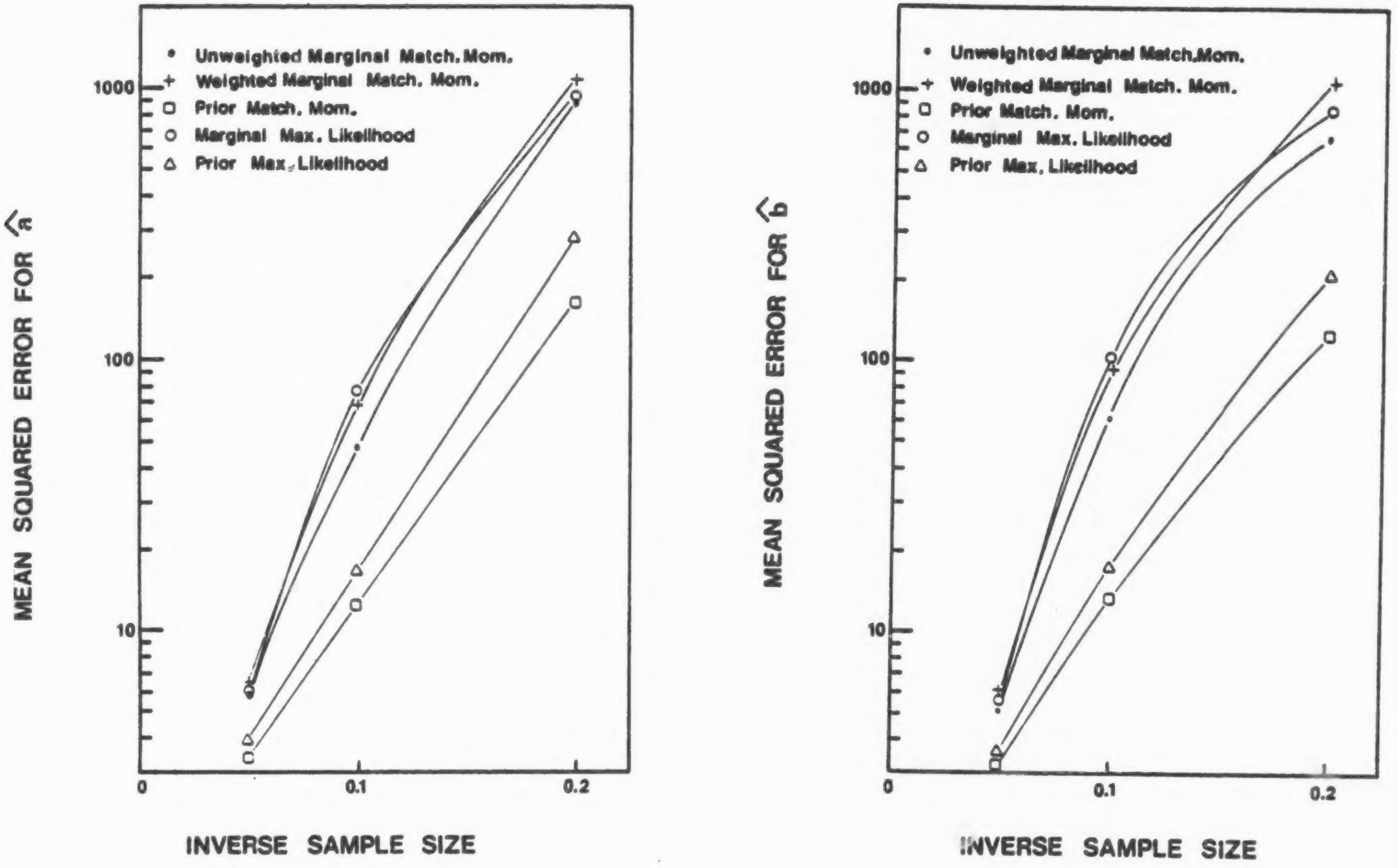


Fig. 4.9 Variation of the mean squared error of the beta parameter estimators with sample size for the symmetric beta distribution (a=b=5).

Table 4.8. The bias and mean squared error of the estimators of the parameters for a symmetric beta prior distribution (a=b=5) as calculated by different estimation techniques from simulated failure data of various sample sizes. Each data set consisted of 500 samples.

Sample Size (N)	WMMM				Prior Matching Moments			
	$\hat{a}-a$	$\hat{b}-b$	MSE($\hat{a}$)	MSE($\hat{b}$)	$\bar{a}-a$	$\bar{b}-b$	MSE($\hat{a}$)	MSE($\hat{b}$)
5	10.98	10.8	1076.	1092.	3.68	3.38	164.0	124.1
10	2.50	2.56	69.1	94.9	0.535	0.533	12.3	13.2
20	0.79	0.764	6.36	5.91	0.110	-0.13	3.47	3.19

Sample Size (N)	Marginal Maximum Likelihood				Prior Maximum Likelihood			
	$\bar{a}-a$	$\bar{b}-b$	MSE($\hat{a}$)	MSE($\hat{b}$)	$\bar{a}-a$	$\bar{b}-b$	MSE($\hat{a}$)	MSE($\hat{b}$)
5	10.3	9.99	936.	862	6.16	5.80	272.	210.
10	2.65	2.70	75.3	102.	1.3	1.30	16.7	12.8
20	0.827	0.805	6.19	5.74	0.208	0.186	3.89	3.51

Sample Size (N)	UWMMM			
	$\bar{a}-a$	$\bar{b}-b$	MSE($\hat{a}$)	MSE($\hat{b}$)
5	7.53	7.14	901.	702.
10	1.60	1.64	46.6	63.1
20	0.460	0.436	5.60	5.15

case, all five methods tend to overestimate the prior parameters for small sample size, and only the simplest method, the prior matching moments technique gives a slight negative bias for samples of size greater than about N=15. Also, as was seen with the skewed case, the two estimation techniques based on the marginal distribution give essentially identical results which are considerably poorer than those obtained with the prior based methods. Thus the prior matching moments techniques had a performance which was as good or better than the other techniques in this symmetric case also.

4.2 Distribution of Estimators for the Mean and Variance of the Prior Distribution

For small sample sizes ($N \leq 20$) all five parameter estimation techniques investigated in this study tended to overestimate values of the parameters a and b for the beta prior distribution. In fact, for very small sample sizes ($N \simeq 5$) and for data generated from the beta prior distribution skewed towards low probability values ($a=1.2$, $b=23$), occasional estimates of a and b were obtained from the marginal-based techniques which were several orders of magnitude too large.

As previously stated, it was observed that whenever an inordinately large value of one beta parameter was obtained, the estimate for the other parameter was also very large. For these overestimation cases, it was observed that a reasonable estimate of the mean of the beta prior was obtained even with these large parameter estimates, since the mean depends only on the ratio a/b, i.e., from Eq. (1.5)

$$\mu = (1 + b/a)^{-1} . \qquad (4.4)$$

The empirical distributions of the estimate of the prior mean was calculated for different sample sizes, by using the estimators $\hat{a}$ and $\hat{b}$ in Eq. (4.4) previously obtained with the simulated failure data for the skewed prior case (true mean = $(1 + 23/1.2)^{-1} = 0.0496$). These distributions are shown in Figs. 4.10-4.13, and the mean and variance of these distributions are given in Table 4.9. Because of the inability of the prior maximum likelihood method to treat low failure probability cases, this method was not included in the analysis.

From these distributions of mean estimators it is seen that no apparent outliers are present. Further the mean of the distributions are all within a small percentage of the true value, although a very slight bias to overestimate the mean is noted. As would be expected, the variances of the distributions decrease as the sample size increases. The most important feature, however, of these distributions of $\hat{\mu}$ is that all four estimation techniques appear to give nearly the same distribution for a given sample size.

Although the presence of outlier estimators for a and b does not affect the distribution of the mean estimators, the high a and b estimates will have a profound effect on the estimation of the variance of the beta prior distribution. The variance of the beta prior is given by (Eq. (1.6))

$$\sigma^2 = [(1 + b/a)(1 + a/b)(a + b + 1)]^{-1} \qquad (4.5)$$

which becomes very small as a and b both become large. Thus the use of outlier estimators $\hat{a}$ and $\hat{b}$ to produce an estimate of the variance for the beta prior will give unrealistically small values. In Figs. 4.14-

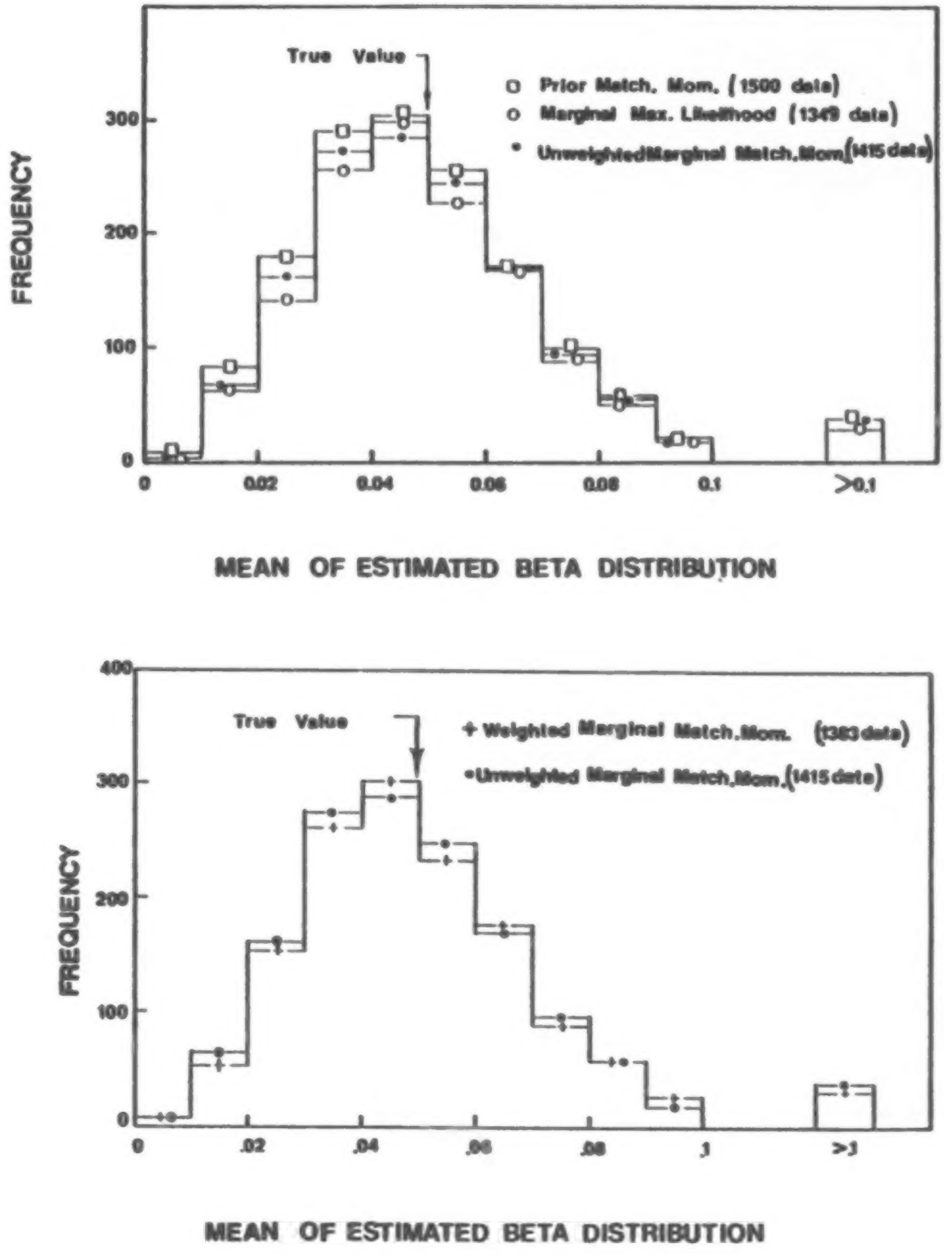


Fig. 4.10 Distribution of the means of the estimated beta prior distributions from samples of size N=5. A true prior mean is 0.0496.

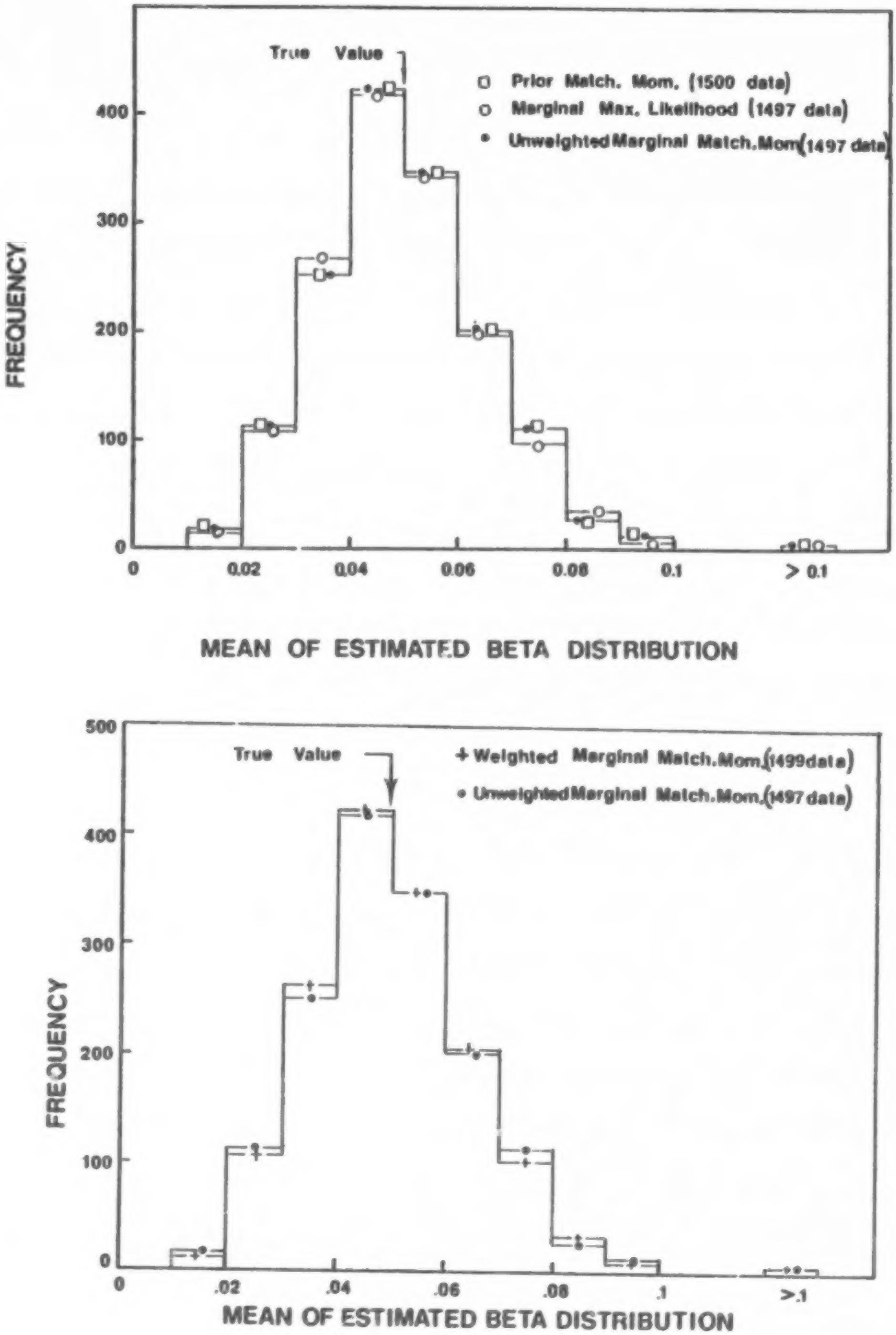


Fig. 4.11 Distribution of the means of the estimated beta prior distributions from samples of size N=10. A true prior mean is 0.0496.

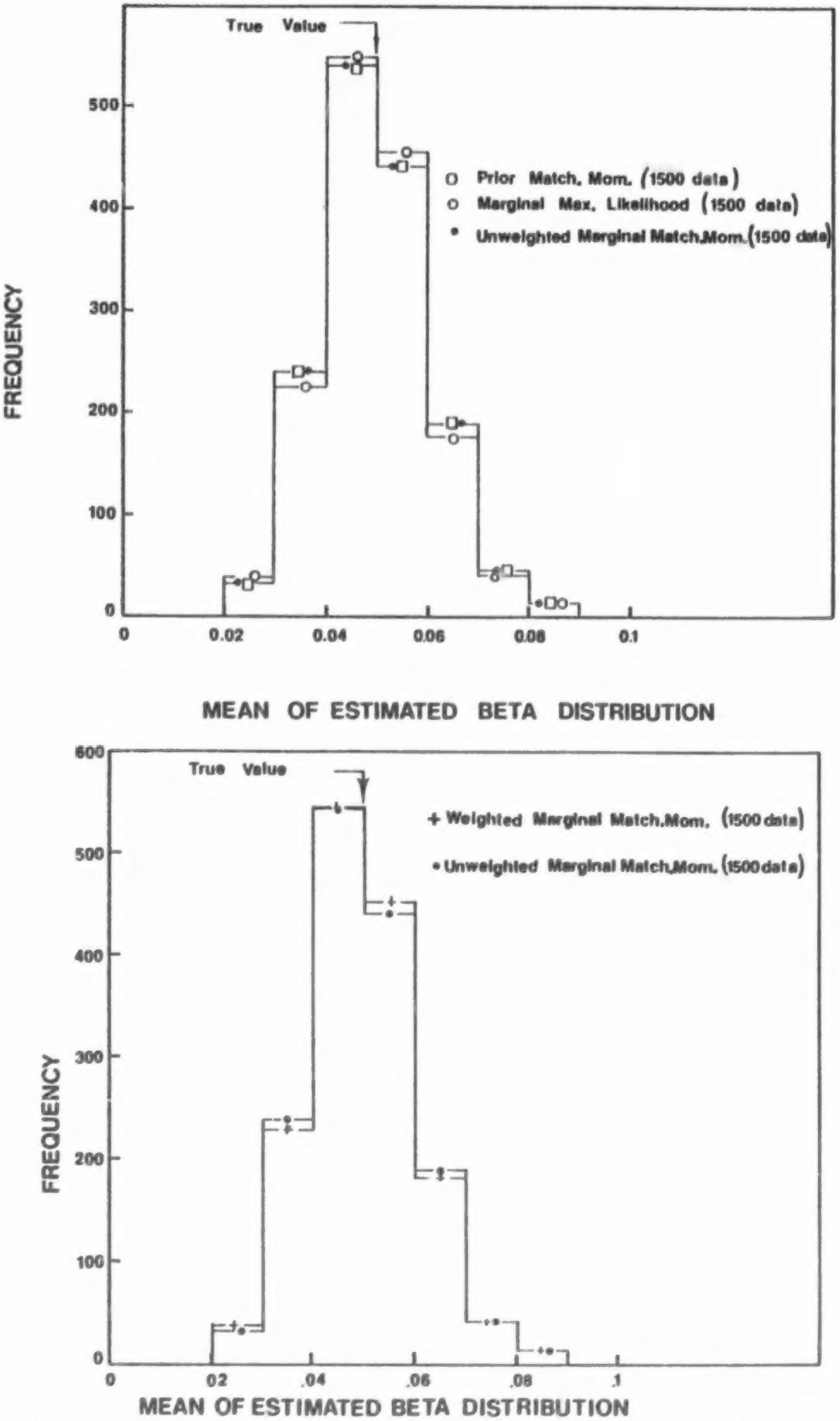


Fig. 4.12 Distribution of the means of the estimated beta prior distributions from samples of size N=20. A true prior mean is 0.0496.

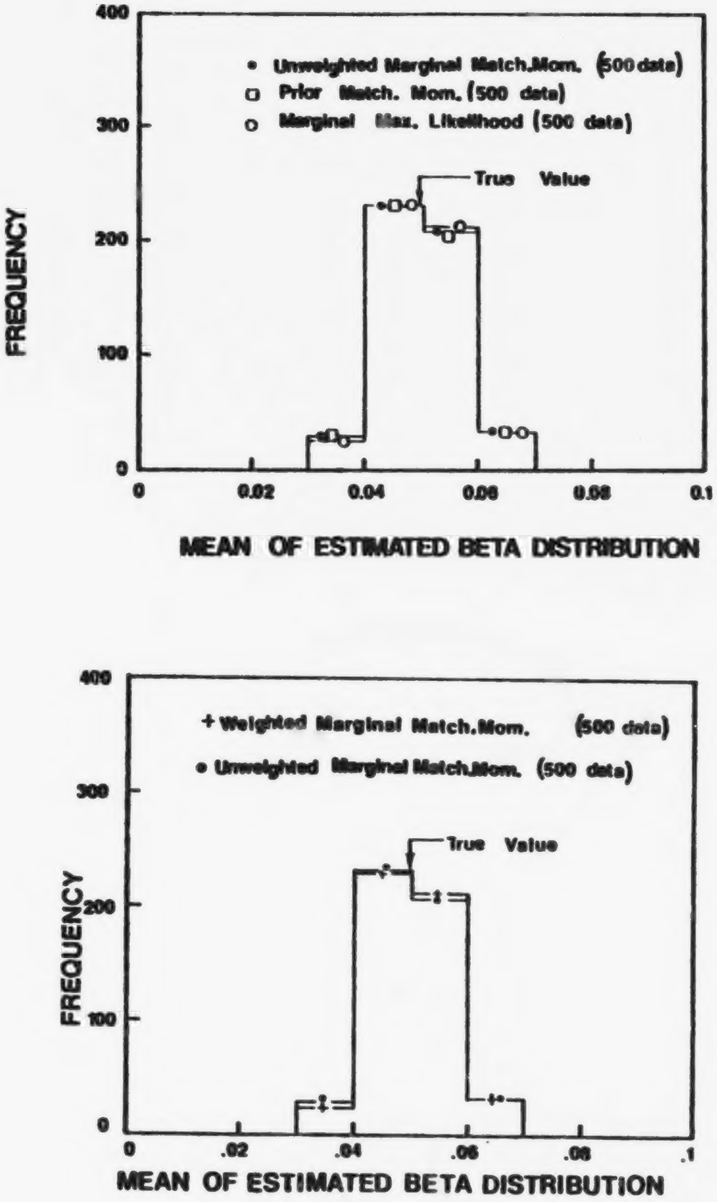


Fig. 4.13 Distribution of the means of the estimated beta prior distributions from samples of size N=50. A true prior mean is 0.0496

Table 4.9. Mean and variance of the estimators for the mean of the beta prior (a=1.2, b=23) for different sample sizes. True prior mean is 0.0496.

Sample	Prior Match. Mom.		Marg. Max. Like.	
Size*	Mean	Variance	Mean	Variance
5	0.0488	0.000423	0.0497	0.000415
10	0.0500	0.000221	0.0498	0.000218
20	0.0496	0.000114	0.0495	0.000112
50	0.0499	0.0000419	0.0499	0.0000419

Sample	WMMM		UWMMM	
Size*	Mean	Variance	Mean	Variance
5	0.0500	0.000416	0.0496	0.000417
10	0.0500	0.000218	0.0500	0.000221
20	0.0496	0.000113	0.0496	0.000114
50	0.0500	0.0000422	0.0499	0.0000419

*1500 samples were used for size 5-20 results; 500 samples were used for size 50 results.

4.17, the distributions of the variance estimators for the prior beta are shown for different sample sizes.

Notice that for small sample sizes (e.g., Fig. 4.4) for which outlier values are expected for the marginal-based estimation methods, the empirical frequency distributions of the variance estimators (Eq. (4.5)) are peaked towards the low end. However as the sample size increases, outlier values for a and b are no longer obtained, and the variance estimator distribution becomes increasing centered around the true variance of $\sigma^2 = 0.00187$. Finally it should be noted from these variance distributions that the distribution produced by the prior matching moments results is always slightly more skewed towards the high values as compared to the distributions for the three marginal-based methods.

In Table 4.10 the mean and variance of these variance estimator distributions are given. It is noted that the mean of the distribution is always slightly less than the true prior variance ($\sigma^2 = 0.00187$) but approaches the true value as the sample size increases. The means of the prior matching moments distributions, however, always overestimate the true mean. More importantly, these overestimates do not appear to approach the true value even as the sample size increases, but rather appear to remain about 20% higher than the true value.

4.3 Distribution of 95-th Percentile Estimators

Of considerable interest in safety analysis is the estimation of the prior distribution at high failure probabilities. One widely used measure of the high probability tail is the 95-th percentile,

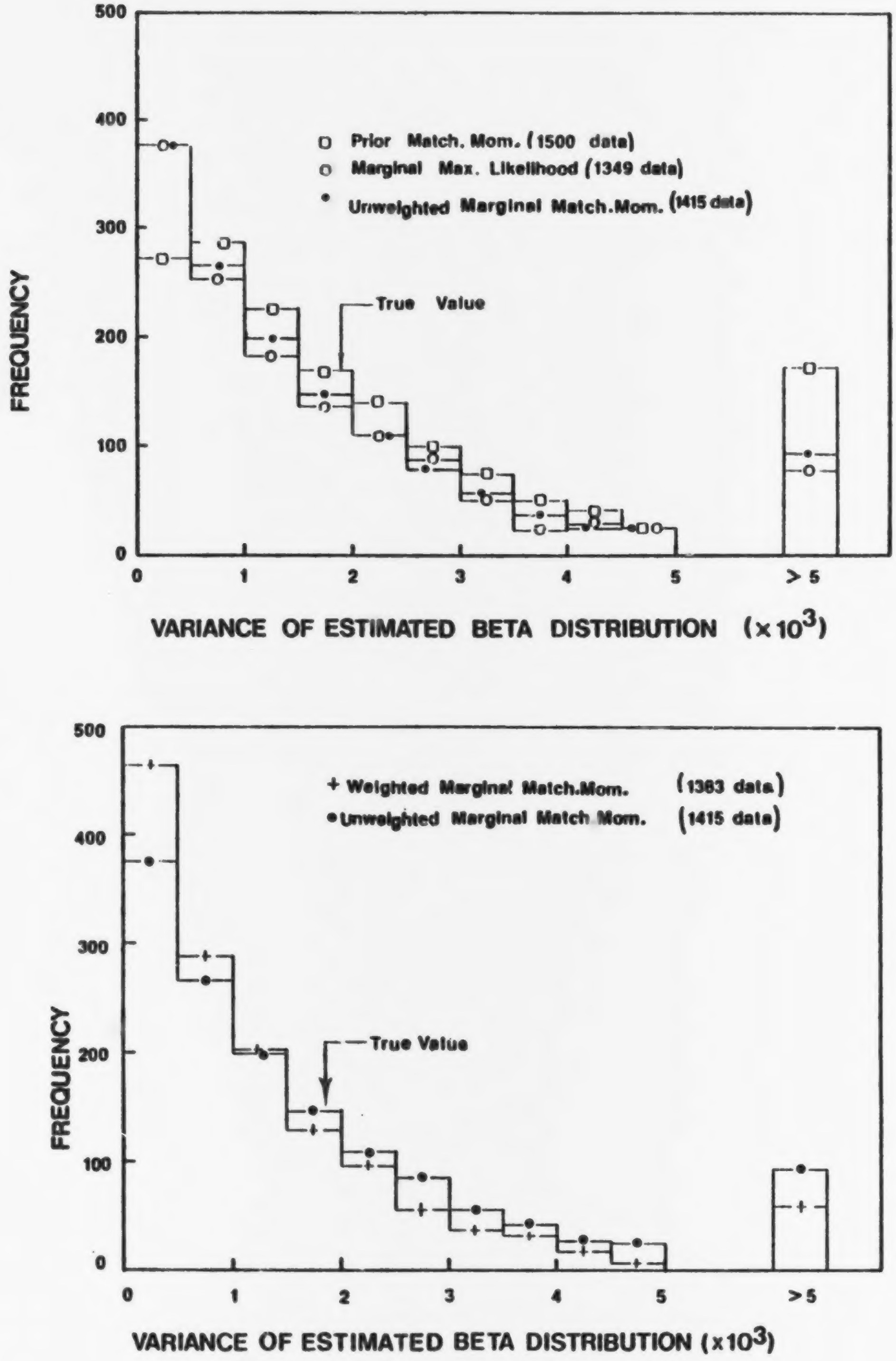


Fig. 4.14 Distribution of the variances of the estimated beta prior distributions from samples of size N=5. A true valúe is 0.00187.

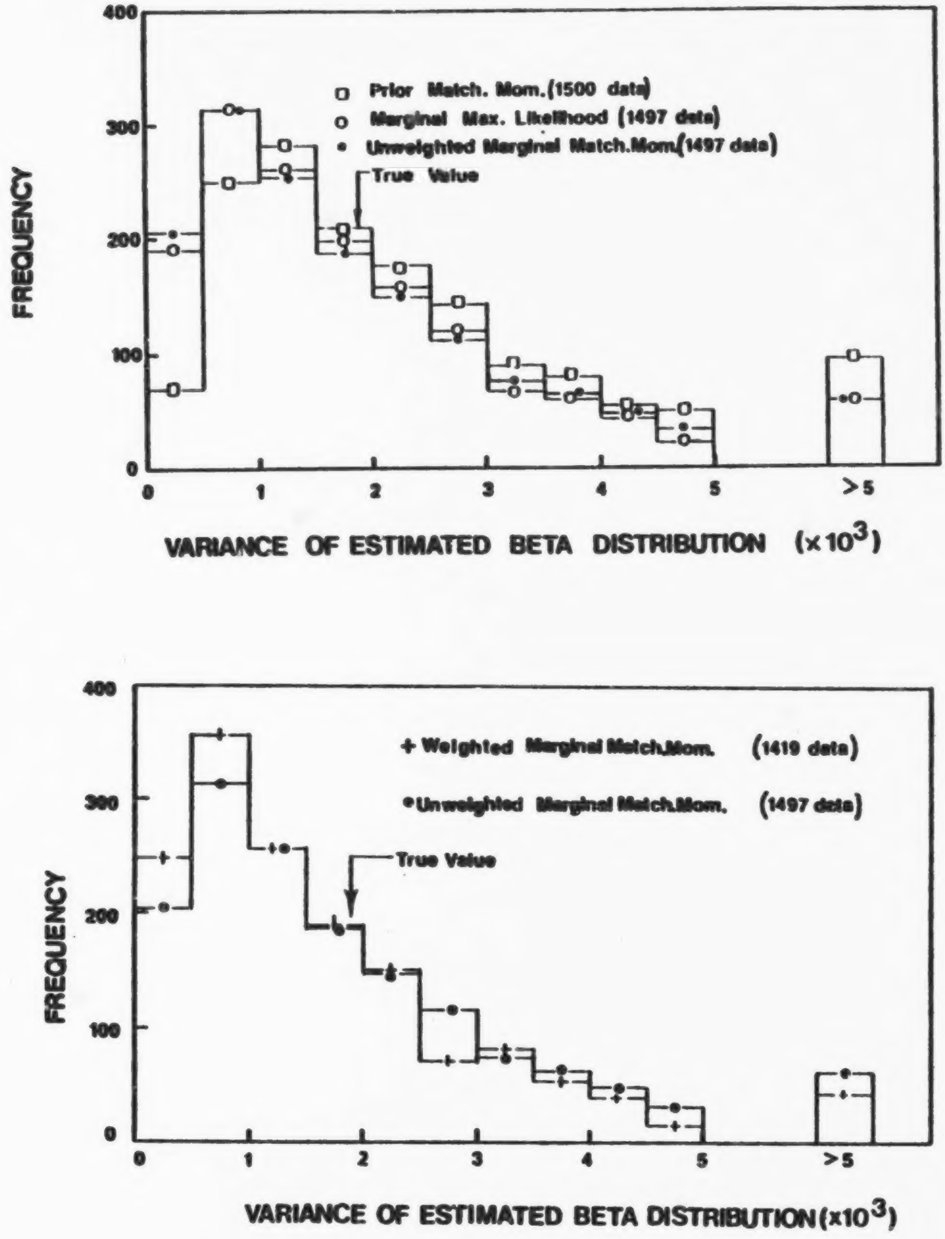
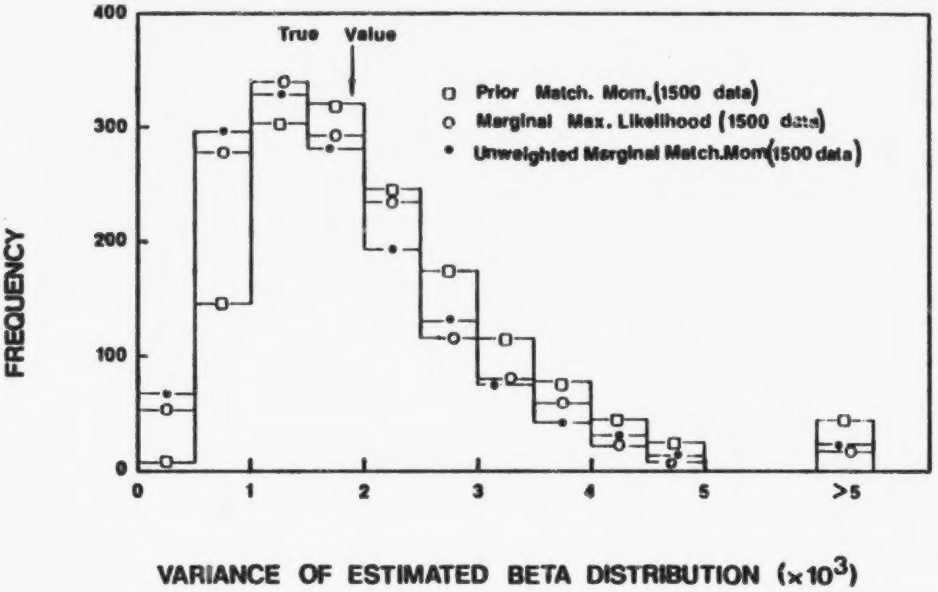


Fig.4.15 Distribution of the variances of the estimated beta prior distributions from samples of size N=10. A true value is 0.00187.



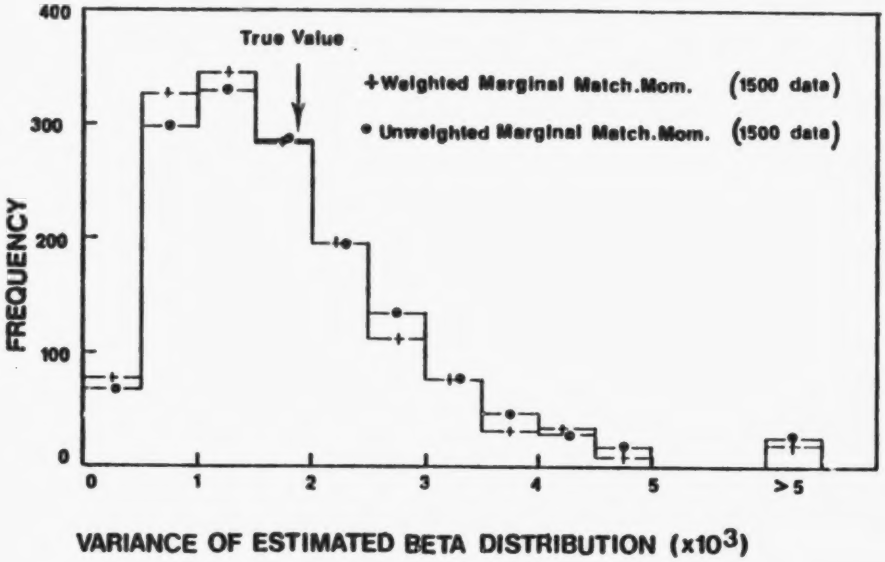


Fig. 4.16 Distribution of the variances of the estimated beta prior distributions from samples of size N=20. A true value is 0.00187.

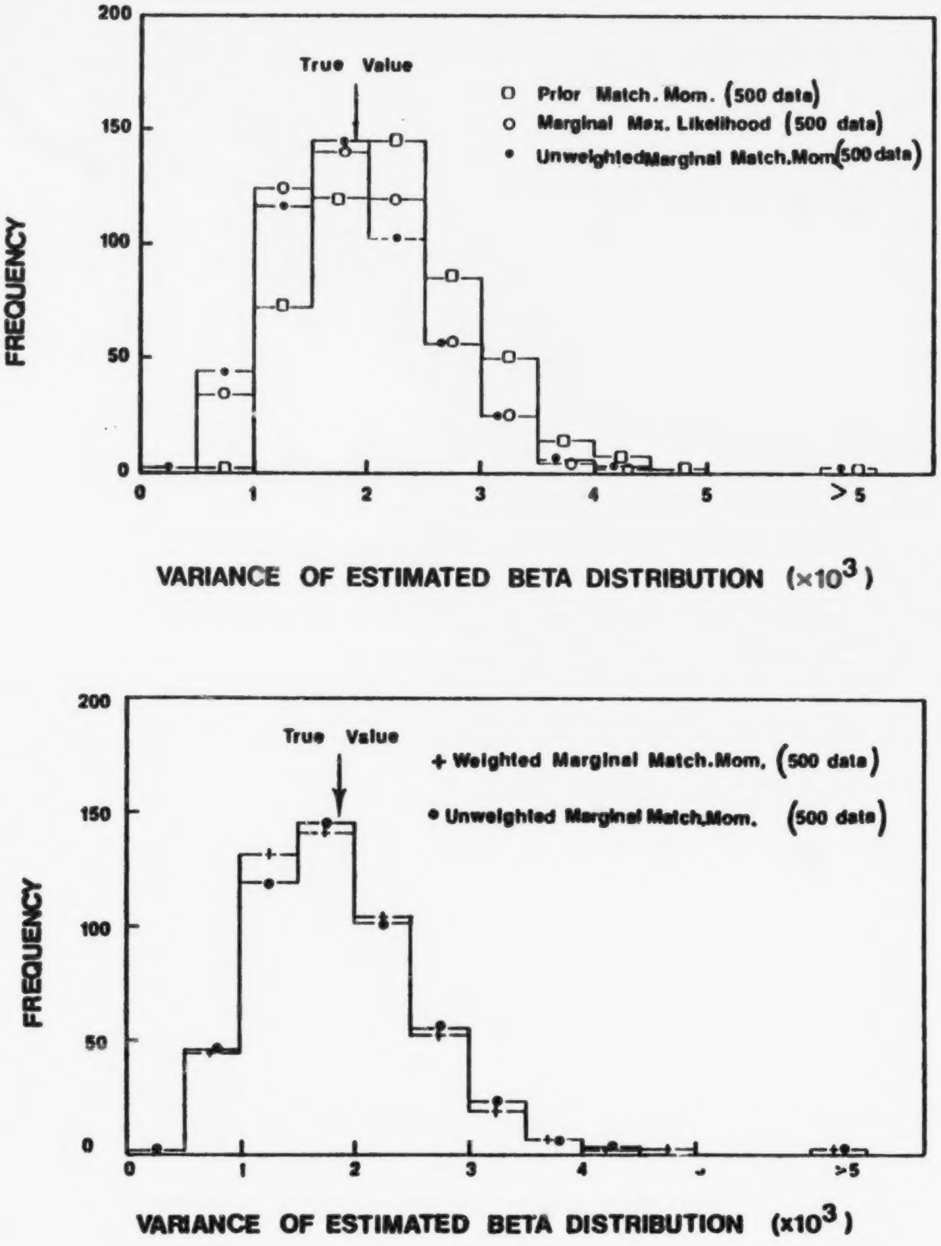


Fig. 4.17 Distribution of the variances of the estimated beta prior distributions from samples of size N=50. A true value is 0.00187.

Table 4.10. Mean and variance of the estimators for the variance of the beta prior (a=1.2, b=23) for different sample sizes. True prior variance is 0.00187.

Sample Size*	Prior Match. Mom. Mean	Var. $[\times 10^5]$	Marg. Max. Like. Mean	Var. $[\times 10^5]$
5	0.00207	0.507	0.00171	0.393
10	0.00227	0.295	0.00185	0.225
20	0.00222	0.145	0.00181	0.102
50	0.00227	0.0558	0.00188	0.0406

Sample Size*	WMMM Mean	Var. $[\times 10^5]$	UWMMM Mean	Var. $[\times 10^5]$
5	0.00141	0.298	0.00180	0.478
10	0.00167	0.215	0.00189	0.275
20	0.00172	0.116	0.00183	0.134
50	0.00184	0.0468	0.00188	0.0517

*1500 samples were used for sizes 5-20 results; 500 samples were used for size 50 results.

i.e., the failure probability, p_{95}, above which there is only a 5% chance that the true failure probability lies for a component described by the prior distribution, g(p). For the beta prior distribution used in this study, the 95-th percentile, p_{95}, is the solution of the following equation:

$$0.95 = \int_0^{p_{95}} g(p)dp = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^{p_{95}} p^{a-1} (1-p)^{b-1} dp. \qquad (4.6)$$

Or equivalently

$$0.95 = I_{p_{95}} (a,b), \qquad (4.7)$$

where the incomplete beta function is defined by

$$I_x(a,b) \equiv \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^{x} p^{a-1} (1-p)^{b-1} dp. \qquad (4.8)$$

Equation (4.7) is readily solved by standard numerical root finding techniques. In this analysis, subroutine RTMI [16] which solves for root of a non-linear equation by Muller's method was used to find the root of Eq. (4.7).

For each simulated failure data set generated for the beta prior which was skewed towards the low probability end (a=1.2, b=23), an estimator of the 95-th percentile was obtained by using the estimators $\hat{a}$ and $\hat{b}$ for each set in Eq. (4.7) and solving numerically for the 95-th percentile. The distribution of the 95-th percentile estimators so obtained are shown in Figs. 4.18-4.21 for the four estimation techniques suitable for analyzing low probability failure data. The mean, variance and median of these distributions are presented in Table 4.11.

From a safety viewpoint, one would like to use an estimation technique which has a low inherent probability of yielding 95-th per-

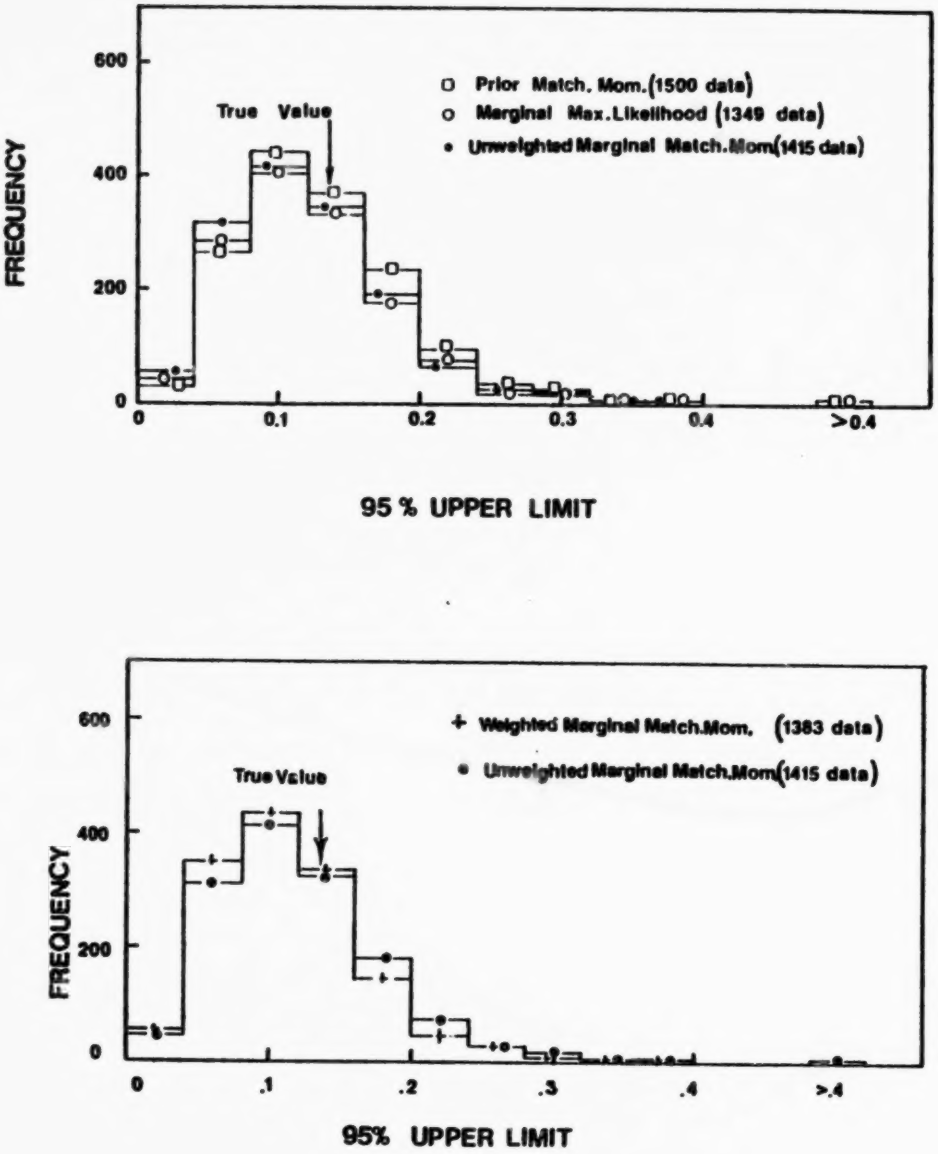


Fig. 4.18 Distribution of 95-th percentiles of the estimated beta prior distributions for samples of size N=5. A true value is 0.136.

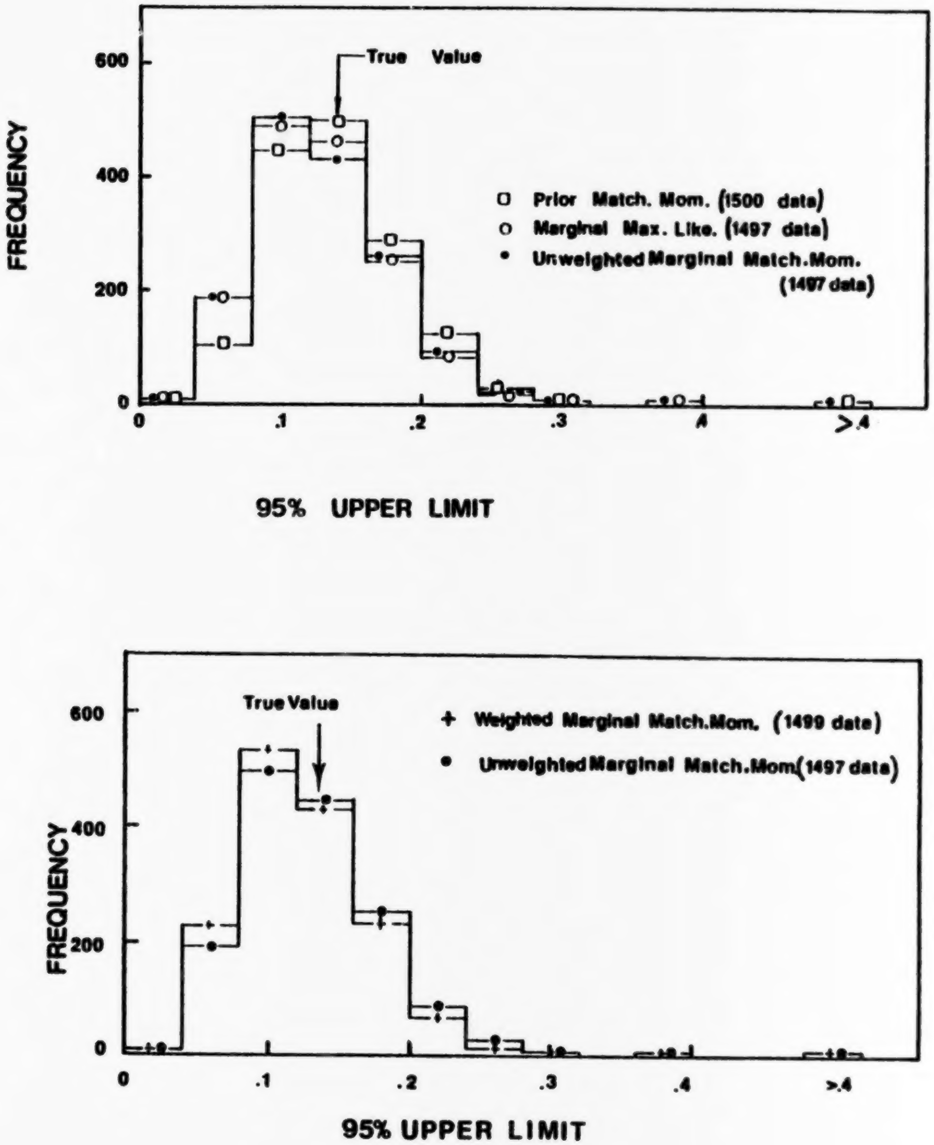


Fig. 4.19 Distribution of the 95-th percentiles of the estimated beta prior distributions for samples of size N=10.
A true value is 0.136.

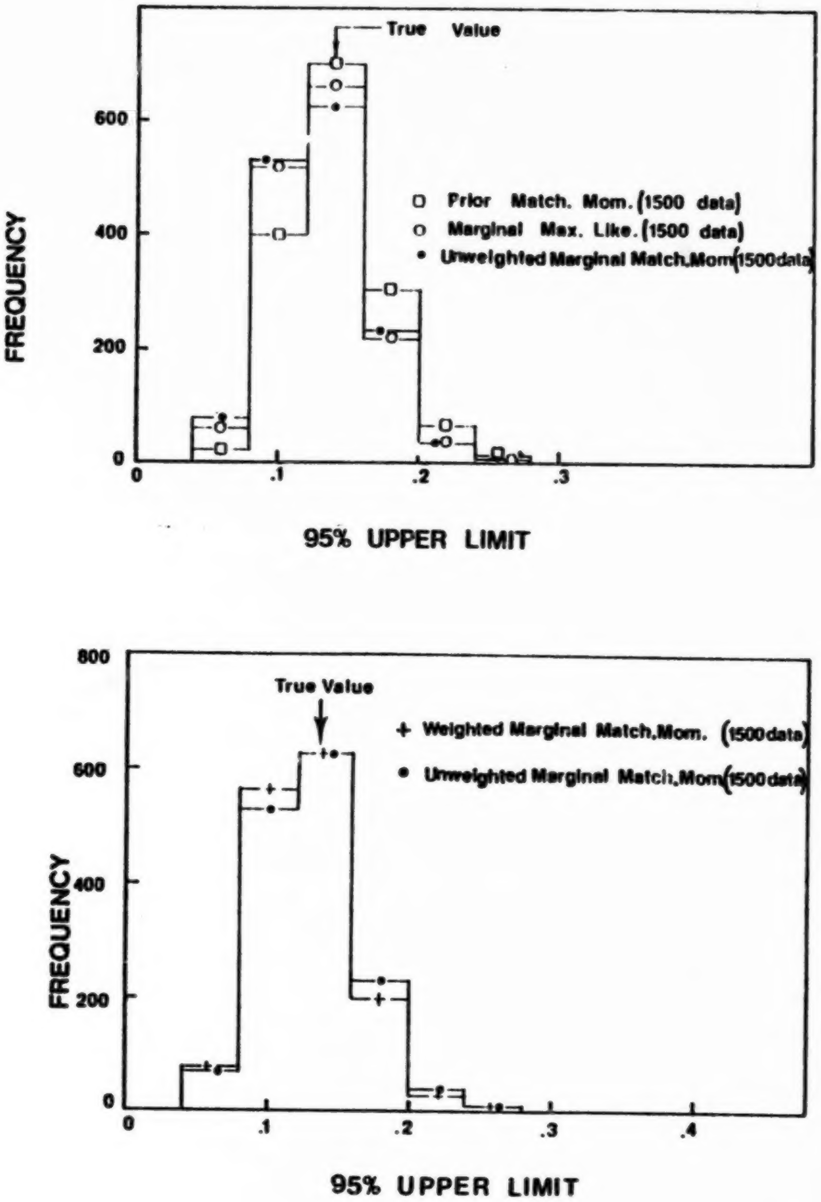
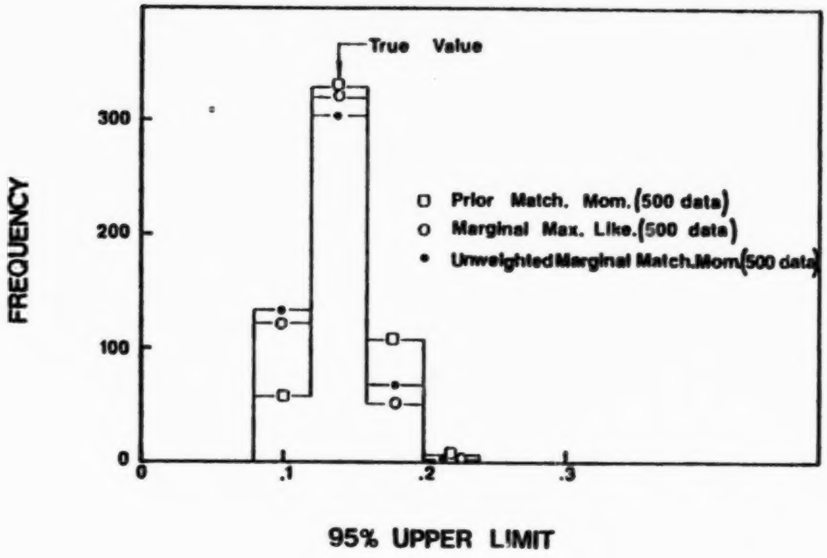


Fig. 4.20 Distribution of 95-th percentiles of the estimated beta prior distributions for samples of size N=20.
A true value is 0.136.



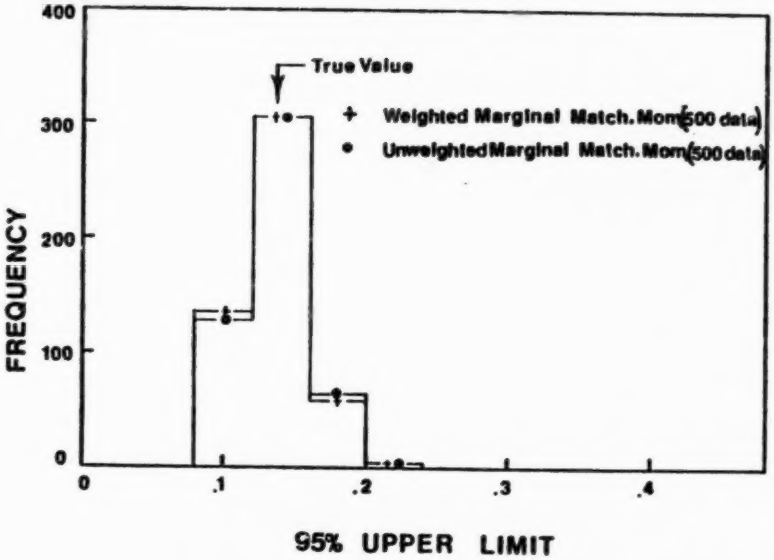


Fig. 4.21 Distribution of the 95-th percentiles of the estimated beta prior distributions for samples of size N=50. A true value is 0.136.

Table 4.11. Median, mean and variance of the distributions of the 95-th percentile estimators. True 95-th percentile = 0.13586.

Sample Size*	Prior Matching Moments			Marginal Max. Likelihood		
	Median	Mean	Var.	Median	Mean	Var.
5	0.121	0.130	0.0035	0.113	0.121	0.0032
10	0.136	0.140	0.0021	0.125	0.129	0.0020
20	0.138	0.141	0.0011	0.129	0.131	0.0010
50	0.144	0.145	0.00044	0.134	0.135	0.00042

Sample Size*	WMMM			UWMMM		
	Median	Mean	Var.	Median	Mean	Var.
5	0.106	0.114	0.0029	0.115	0.123	0.0035
10	0.119	0.124	0.0020	0.125	0.130	0.0022
20	0.125	0.128	0.0011	0.128	0.131	0.0012
50	0.133	0.134	0.00045	0.134	0.135	0.00047

*1500 samples were used for size 5-20; 500 samples for size 50.

centile estimates which are very much less than the true value. In other words, if the estimator is biased, then it would be better if it were biased so as to yield overestimates of p_{95} (with hopefully small minimum mean square error). Further, there should be little if any chance of yielding outliers or values of $\hat{p}_{95}$ which are orders of magnitude less than the true value. For the present case the true value of the 95-th percentile for a=1.2 and b=23 is $p_{95}=0.13586$. In Table 4.12, the number of simulated data samples which yielded estimators greater than or less than the true p_{95} are given. Notice that for small samples all four estimation methods are non-conservative ($\text{Prob}\{\hat{p}_{95}<p_{95}\}>0.5$), while as the sample size increases, the prior matching moments becomes increasingly conservative while the medians for the other three methods approach the true p_{95} value.

From Table 4.11, all three methods are seen to yield distributions for $\hat{p}_{95}$ with almost equal variance. However, the three marginal-based estimation techniques yield distributions with means and medians smaller than the true value for all sample sizes although as the sample size increases the medians and means increase and approach the true value of p_{95}. The simple prior matching moments technique also yields distributions of $\hat{p}_{95}$ whose mean and median also increase with increasing sample size, but unlike the other techniques, for sample sizes greater than about seven, the means and medians become greater than the true values, i.e., the distribution becomes conservative. Further for very large sample sizes this positive bias does not disappear, although the bias may not be significantly large.

Table 4.12. Number and percent of simulated failure data samples which yielded estimated 95-th percentiles greater than (GT) or less than (LT) the true value of 0.13586.

Sample Size	Prior Match. Mom.				Marg. Max. Likelihood			
	LT		GT		LT		GT	
	No.	%	No.	%	No.	%	No.	%
5	890	59.3	610	40.7	873	74.7	476	35.3
10	755	50.3	745	49.7	883	59.0	614	41.0
20	701	46.7	799	53.3	820	54.7	680	45.3
50	176	35.2	324	64.8	261	52.2	239	47.8

Sample Size	WMMM				UWMMM			
	LT		GT		LT		GT	
	No.	%	No.	%	No.	%	No.	%
5	978	70.7	405	29.3	908	64.2	507	35.8
10	953	63.6	546	36.4	882	58.9	615	41.1
20	873	58.2	627	41.8	893	59.5	607	40.5
50	277	55.4	223	44.6	270	54.0	230	46.0

For small sample sizes (N=5) (see Fig. 4.18) all four methods yield some estimators $\hat{p}_{95}$ in the lowest value bin (0-0.04). These values are, of course, not conservative. Of considerable concern is how these low estimates are distributed in this low end bin. Since the marginal-based estimation techniques occassionally yield very large estimators for a and b, i.e., outliers, the resulting estimated prior distribution will have a very small variance and hence the 95-th percentile will be only slightly greater than the mean. If the mean should turn out to be very small, the $\hat{p}_{95}$ values for these outliers could be very much smaller than the true value. Clearly such a feature of these estimation techniques would preclude their use in safety analyses. In Table 4.13, the lowest 5 values of $\hat{p}_{95}$ found in the present simulation study are listed. It is seen that only three estimates are smaller than 10% of the true value, and hence the possibility of obtaining in the $\hat{p}_{95}$ distribution severe outliers which are orders of magnitude smaller than the true value does not appear to be very likely.

4.4 Fraction of the Estimated Prior Distribution Above the True 95-th Percentile

The extent of the high probability tail of the estimated beta prior distribution is of considerable concern in safety analysis. In the previous section the distribution of the 95-th percentiles of the estimated prior distributions was discussed. An alternative perspective is to consider the fraction of the estimated prior that is supported above the true 95-th percentile, i.e., the probability that the estimated failure probability is greater than the true 95-th percentile. This quantity is given by

Table 4.13. Smallest 95-th percentile estimators observed for simulated failure data samples of size N. True value of the 95-th percentile is 0.13586.

Weighted Marg. Matching Moments

N=5	N=10	N=20	N=50
0.0193	0.0362	0.0428	0.0863
0.0206	0.0364	0.0446	0.0871
0.0221	0.0371	0.0503	0.0881
0.0223	0.0387	0.0533	0.0845
0.0234	0.0395	0.0554	0.0922

Prior Matching Moments

N=5	N=10	N=20	N=50
0.0115	0.0385	0.0592	0.0974
0.0196	0.0451	0.0622	0.101
0.0242	0.0491	0.0658	0.101
0.0243	0.0500	0.0673	0.101
0.0256	0.0509	0.0695	0.102

Marginal Maximum Likelihood

N=5	N=10	N=20	N=50
0.0152	0.0269	0.0426	0.0848
0.0154	0.0306	0.0461	0.0870
0.0170	0.0360	0.0503	0.0892
0.0209	0.0369	0.0505	0.0922
0.0239	0.0400	0.0572	0.0924

Unweighted Marginal Matching Moments

N=5	N=10	N=20	N=50
0.0124	0.0291	0.0417	0.0870
0.0131	0.0355	0.0498	0.0877
0.0218	0.0360	0.0544	0.0893
0.0227	0.0389	0.0563	0.0903
0.0235	0.0392	0.0564	0.0905

$$\text{Prob}\left\{\text{estimated } p \geq p_{95}^{\text{true}}\right\} = \int_{p_{95}^{\text{true}}}^{1} g_{\text{est}}(p)\, dp\ , \qquad (4.9)$$

where p_{95}^{true} is the 95-th percentile of the beta distribution used to generate the simulated failure data (a=1.2, b=23), and $g_{\text{est}}(p)$ is the estimated prior distribution for a particular failure data sample (i.e., a beta distribution with $a=\hat{a}$ and $b=\hat{b}$). Equivalently Eq. (4.9) can be written in terms of the incomplete beta function as

$$\text{Prob}\left\{\text{estimated } p \geq p_{95}^{\text{true}}\right\} = 1 - I_{p_{95}}^{\text{true}}\,(\hat{a},\hat{b}) \qquad (4.10)$$

where $I_{p_{95}}^{\text{true}}(\hat{a},\hat{b})$ is incomplete beta function defined in Eq. (4.8).

If the estimation technique used to analyze the failure data should yield estimators $\hat{a}$ and $\hat{b}$ equal to the true values of the beta prior, then the probability given by Eq. (4.9) would equal 0.05. Of course, the estimation techniques will not in general yield exact values for the beta parameters, and those methods which tend to yield estimated priors skewed more towards higher probability values than the true prior are preferred for safety analysis since the resulting estimated failure probabilities will be overestimated and hence conservative.

The distribution of the probability estimates given by Eq. (4.9), for the four parameter estimation techniques suitable for analyzing low failure probability data, are shown in Figs. 4.22-4.25. It is seen that all four estimation methods yield a considerable portion of values of $\text{Prob}\{p \geq p_{95}^{\text{true}}\}$ below the ideal value of 0.05. As the sample size increases, these distributions become increasingly centered about 0.05. However, the distribution for $N \leq 20$ are all highly skewed towards small probabilities with a long slowly decaying behavior at high values. The

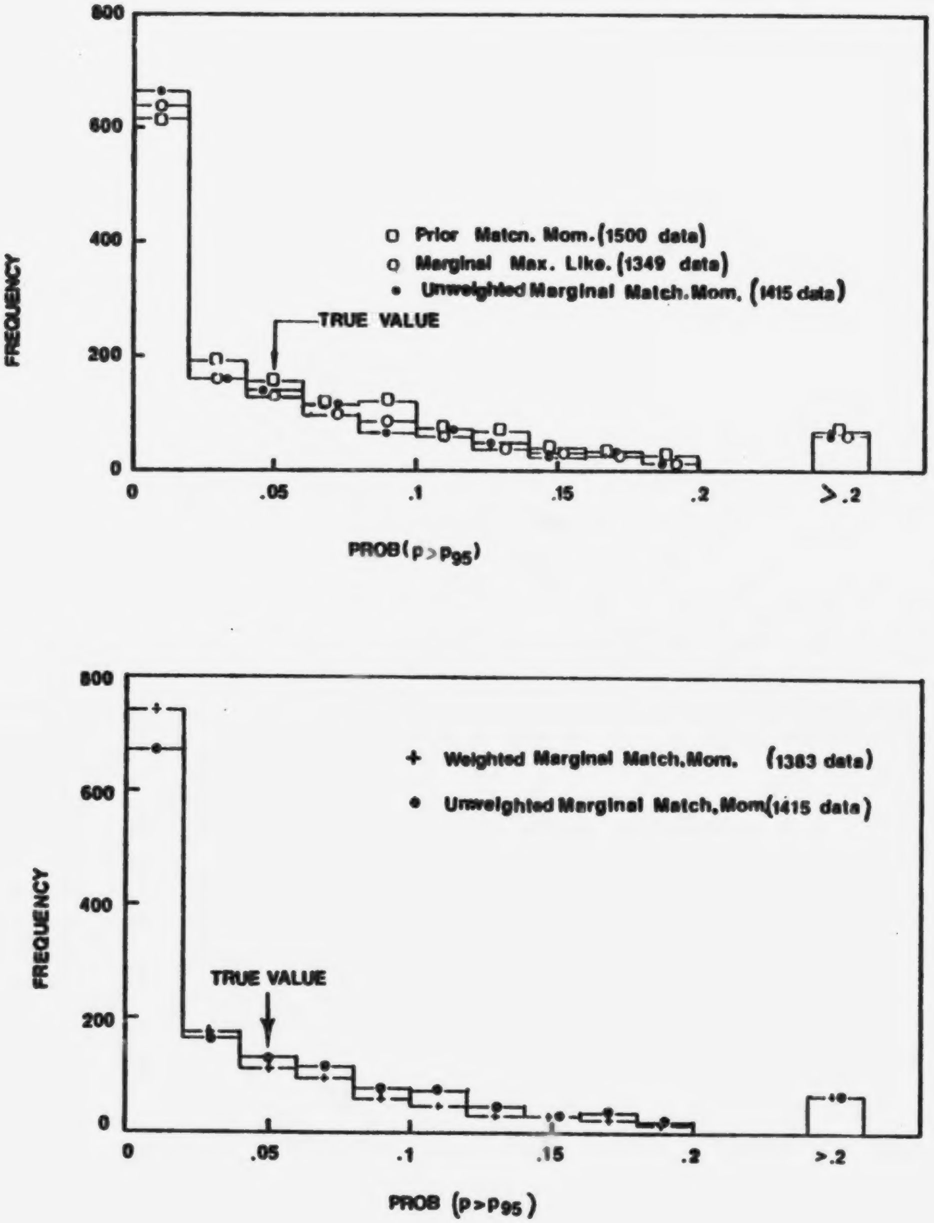


Fig. 4.22 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data(a=1.2,b=23). Size of samples used to obtain estimates was N=5.

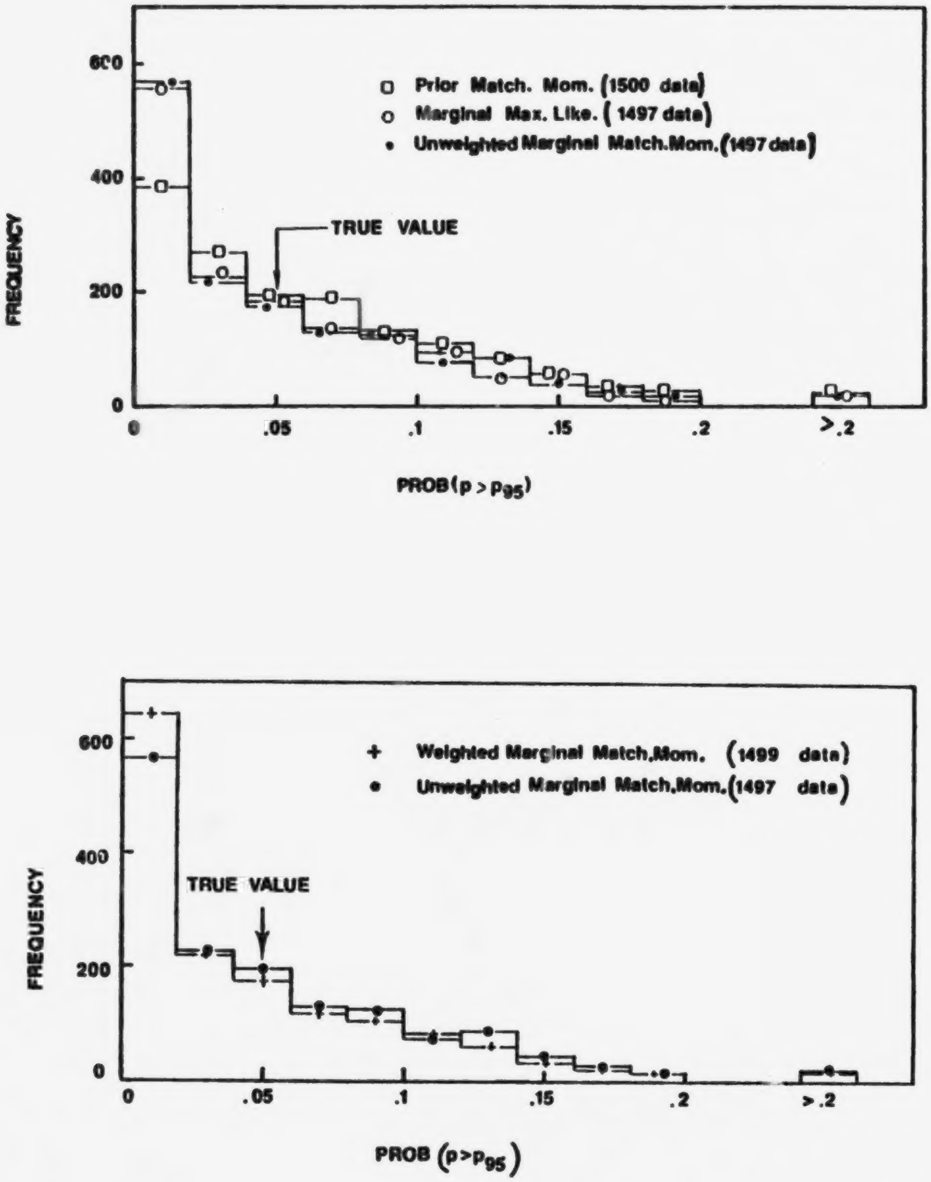


Fig. 4.23 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data(a=1.2,b=23). Size of samples used to obtain estimates was N=10.

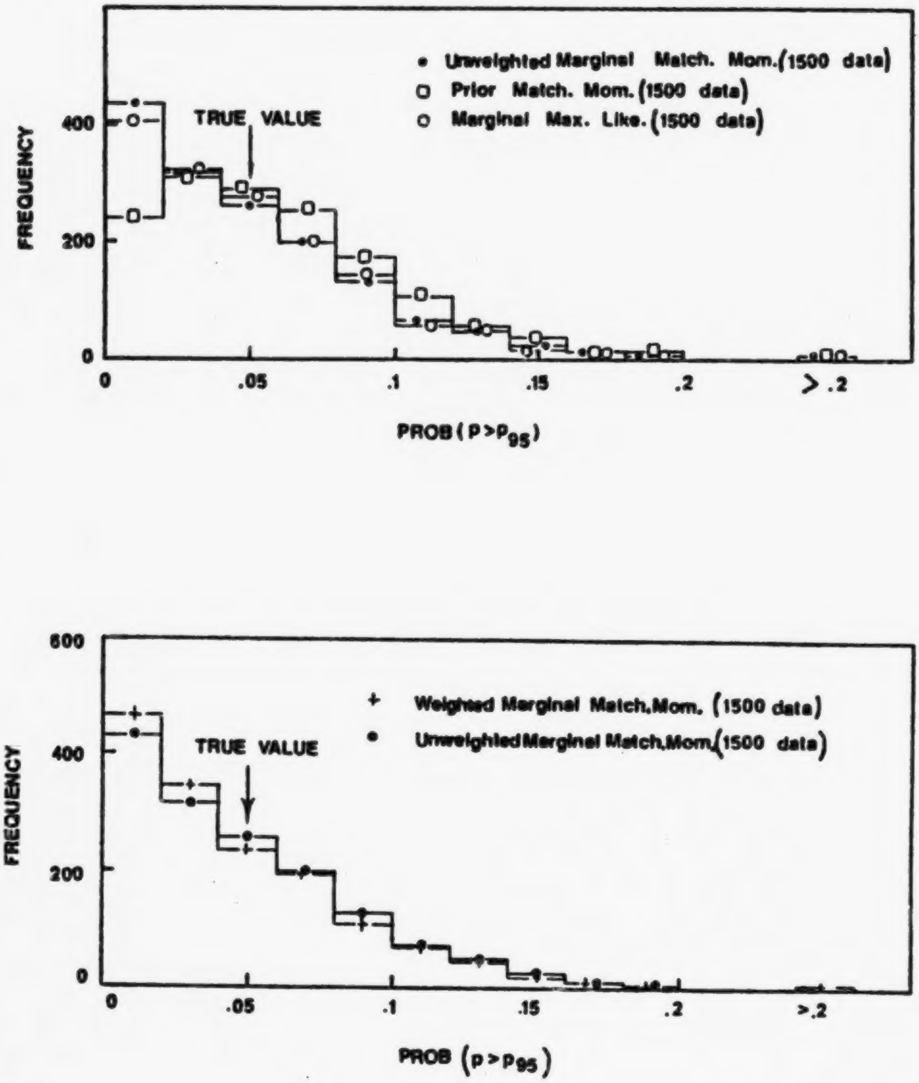


Fig. 4.24 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data(a=1.2,b=23). Size of samples used to obtain estimates was N=20.

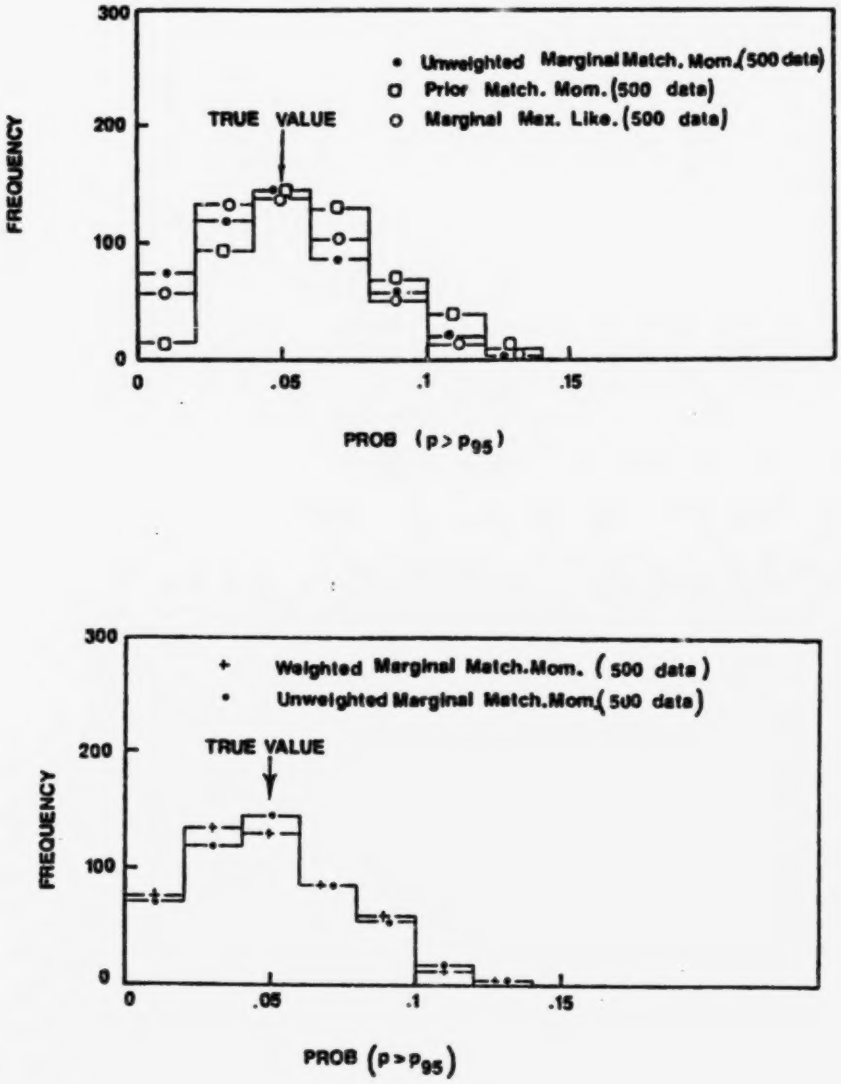


Fig. 4.25 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data(a=1.2,b=23). Size of samples used to obtain estimates was N=50.

prior matching moments method in all cases appears to be slightly more conservative by giving a distribution which is not as concentrated at the low probability values as compared to the distributions obtained with the other three estimation techniques.

The median, mean and variance of these distributions are presented in Table 4.14. From these results the variances for all four methods are within a few percent of each other although the mean for the prior matching moment distribution is considerably higher than that for the distributions produced by the marginal-based methods. Moreover, even for large sample sizes the mean of the distribution for the prior matching moments method is about 20% greater than the ideal value of 0.05. The marginal-based methods, in contrast, appear to approach the ideal value as the sample size becomes sufficiently large.

4.5 Comparison of Maximum Likelihood Variance Bounds to Measured Variances

In Section 2.7 expressions for the variances and covariance of the parameter estimators were derived for the marginal maximum likelihood method. Although these expressions are strictly asymptotic values, the expressions are often used as actual estimators of the variance or covariance of the parameter estimates for finite size data samples. Since the values of the variances and covariances of the parameter estimates are important for error propagation one would like to know how close these maximum likelihood estimated values are to the true values of the variances and covariance.

The actual variances and covariance for the parameter estimators found in the simulation study are listed in Table 4.4. Because of the

Table 4.14. Median, mean and variance of the distribution for the Prob$\{p \geq p_{95}^{true}\}$. For samples of size 5, 10, and 20, 1500 simulated failure data sets were used, while for the size 50 sample, 500 sets were used. Beta prior parameters are a=1.2 and b=23.

Sample Size	Prior Match. Mom.			Marg. Max. Like.		
	Median	Mean	Var.	Median	Mean	Var.
5	0.0321	0.0570	0.0045	0.0230	0.0493	0.0043
10	0.0498	0.0616	0.0027	0.0363	0.0511	0.0026
20	0.0532	0.0595	0.0015	0.0415	0.0489	0.0015
50	0.0596	0.0618	0.00065	0.0478	0.0508	0.00065

Sample Size	WMMM			UWMMM		
	Median	Mean	Var.	Median	Mean	Var.
5	0.0142	0.0425	0.0041	0.0242	0.0508	0.0044
10	0.0287	0.0462	0.0025	0.0350	0.0516	0.0027
20	0.0367	0.0456	0.0015	0.0400	0.0484	0.0016
50	0.0467	0.0491	0.00068	0.0473	0.0502	0.00070

presence of estimator outliers for small sample sizes ($N \leq 10$) obtained with both marginal-based estimation techniques, the experimental values of variances and covariance depends greatly on how these outliers are treated. In this study estimators greater than 100 times the true beta parameter values ($a=1.2$, $b=23$) were ignored.

To evaluate the effectiveness of using the maximum likelihood expressions as estimators, simulated failure data samples were selected which produced either excellent or very poor parameter estimates. With these data samples the marginal maximum likelihood variance bounds were calculated from Eqs. (3.54)-(3.56). The results for the "good" and "bad" data samples are shown in Table 4.15 and the data samples themselves are given in Table 4.16. From these results it is seen that the "bad" data samples which yield inordinately large values for $\hat{a}$ and $\hat{b}$, also produce extremely large estimates for the variances and covariance and are much larger than the empirical estimates in Table 4.4.

The maximum likelihood estimates for the "good" data samples appear much more reasonable and are generally smaller than the empirically observed variances listed in Table 4.4. To compare these maximum likelihood estimates to the variances and covariance measured from the distributions of the parameter estimators, the ratio of the measured value to the likelihood bound was calculated. These ratios are presented in Table 4-17 for each of the four estimation techniques suitable for the low failure probability case studied. From these results it is seen that the empirical variances of the parameter estimator as determined by the prior matching moment technique are much closer to the

Table 4.15. Variance bounds [bnd($\hat{a}$) and bnd($\hat{b}$)], and the covariance bounds [bnd($\hat{a}$,$\hat{b}$)] for parameter estimators [$\hat{a}$ and $\hat{b}$], as calculated by the marginal maximum likelihood method for selected simulated failure data samples. True values of the beta parameters is a=1.2 and b=23. The selected data samples are given in Table 4.16.

Sample Size	ID No.	$\hat{a}$	$\hat{b}$	bnd($\hat{a}$)	bnd($\hat{b}$)	bnd($\hat{a}$,$\hat{b}$)
5	1	1.2444	22.823	0.89839	393.129	16.179
	2	528.92	11338.	$3.0843\text{x}10^{8}$	$1.417\text{x}10^{11}$	$6.6111\text{x}10^{9}$
10	3	1.2673	23.541	0.42806	193.50	7.8072
	4	2080.8	40183.	$3.9119\text{x}10^{10}$	$1.4589\text{x}10^{13}$	$7.5545\text{x}10^{11}$
20	5	1.2248	22.720	0.20962	94.534	3.8150
	6	7.1495	137.61	19.074	7309.5	366.41
50	7	1.1728	23.094	0.076788	39.481	1.4846
	8	2.8889	58.522	0.67451	308.08	13.580

Table 4.16. Selected simulated failure data samples used to estimate variance bounds in Table 4.15. Data were simulated from a beta binomial with parameters a=1.2 and b=23. Data are read from left to right with the number of failures, k_i, following the number of tries, n_i.

Sample Size	ID No.	(n_i,k_i)									
5	1	45	4	216	5	213	25	92	0	260	9
	2	246	12	249	13	227	4	167	8	255	14
10	3	100	3	109	9	83	11	242	5	287	19
		247	4	116	6	248	5	195	21	256	0
	4	45	3	265	14	43	1	164	7	288	14
		44	4	180	15	247	13	163	4	247	8
20	5	46	4	43	1	276	35	139	0	168	16
		160	3	84	9	175	2	169	0	219	13
		264	37	271	22	247	12	111	4	106	1
		243	16	111	1	191	9	105	1	228	9
	6	227	4	91	5	287	17	184	3	121	10
		264	26	137	6	286	8	255	9	118	8
		175	7	128	3	31	2	225	12	150	11
		166	3	34	3	150	11	188	10	173	7
50	7	261	20	33	0	281	11	237	29	203	8
		157	35	227	7	44	1	245	6	59	1
		155	8	176	10	48	2	192	14	82	1
		241	7	150	25	255	4	265	3	131	4
		119	14	148	6	102	8	103	5	87	7
		266	0	137	0	178	1	261	34	280	2
		144	4	227	11	284	7	244	6	56	1
		184	3	101	4	196	2	213	3	125	16
		137	0	172	0	122	19	218	8	261	9
		80	7	60	2	254	16	241	5	263	6
	8	209	2	77	2	158	13	168	18	213	1
		209	19	63	0	196	9	30	2	104	1
		224	8	173	11	155	5	143	7	266	20
		250	27	42	0	290	17	153	7	101	4
		286	18	213	15	132	6	56	1	62	4
		68	3	273	14	199	2	116	4	80	3
		142	14	140	9	208	7	243	13	235	19
		287	12	204	0	167	8	300	16	262	8
		226	13	142	6	227	2	169	6	124	6
		165	7	267	3	97	8	163	15	193	1

Table 4.17. Ratio of measured variances and covariances of the parameter estimators (listed in Table 4.5) to the marginal maximum likelihood bounds (bnd) (listed in Table 4.15) for the "good" data samples.

Sample Size	Prior Matching Moments			Marg. Max. Likelihood		
	$\frac{\mathrm{var}(\hat{a})}{\mathrm{bnd}(\hat{a})}$	$\frac{\mathrm{var}(\hat{b})}{\mathrm{bnd}(\hat{b})}$	$\frac{\mathrm{cov}(\hat{a},\hat{b})}{\mathrm{bnd}(\hat{a},\hat{b})}$	$\frac{\mathrm{var}(\hat{a})}{\mathrm{bnd}(\hat{a})}$	$\frac{\mathrm{var}(\hat{b})}{\mathrm{bnd}(\hat{b})}$	$\frac{\mathrm{cov}(\hat{a},\hat{b})}{\mathrm{bnd}(\hat{a},\hat{b})}$
5	4.92	9.64	6.36	61.1	69.7	71.1
10	1.28	1.48	1.29	13.1	21.1	17.6
20	1.01	1.06	0.993	2.72	3.59	3.20
50	0.875	0.773	0.829	1.48	1.48	1.53

Sample Size	WMMM			UWMMM		
	$\frac{\mathrm{var}(\hat{a})}{\mathrm{bnd}(\hat{a})}$	$\frac{\mathrm{var}(\hat{b})}{\mathrm{bnd}(\hat{b})}$	$\frac{\mathrm{cov}(\hat{a},\hat{b})}{\mathrm{bnd}(\hat{a},\hat{b})}$	$\frac{\mathrm{var}(\hat{a})}{\mathrm{bnd}(\hat{a})}$	$\frac{\mathrm{var}(\hat{b})}{\mathrm{bnd}(\hat{b})}$	$\frac{\mathrm{cov}(\hat{a},\hat{b})}{\mathrm{bnd}(\hat{a},\hat{b})}$
5	57.9	63.6	61.2	55.7	70.7	66.1
10	28.7	29.8	32.1	31.5	33.9	36.1
20	3.82	4.57	4.43	4.07	4.56	4.53
50	2.28	2.06	2.32	2.23	2.01	2.27

likelihood estimates than are the variances for the estimators as determined by either of the marginal based techniques. The marginal-based estimators, $\hat{a}$ and $\hat{b}$, have empirical variances which are many times larger than the likelihood expressions for samples less than 20 in size, although the variances still appear to approach the bounds as the sample size becomes very large.

It should be emphasized that the above conclusions hold for particular examples of "good" failure data. Whether they hold true on the average for all data samples is the subject of further investigation. However, it is seen by the "bad" data samples used here, that the likelihood bounds are capable of yielding completely unrealistic values, and hence for the analysis of a single failure data sample, care must be used in using the likelihood bounds as estimates for the variances of the prior parameter estimators.

5.0 CONCLUSIONS

Based on simulated failure data which were generated from both skewed and symmetric beta-binomial marginal distributions, it has been found that the beta prior parameter estimators $\hat{a}$ and $\hat{b}$, obtained from the prior matching moments estimation techniques possess the least bias and mean-squared error among the five estimation methods investigated for small sample sizes ($N<50$). Moreover, only this method has closed-form expressions for the parameter estimators and always yields realistic prior parameter estimators for all simulated data sample of various sizes. The results also showed that the prior maximum likelihood method is not suitable for low failure probability case since this method is infeasible for any failure data sample for which a single component with zero failures is observed. The two most complicated methods, the weighted marginal matching moments and marginal maximum likelihood, had the biggest biased and mean-squared errored estimators. In addition, these two techniques would occasionally fail to yield parameter estimators or yield outlier estimators which were much too large in size. This deficiency was more severe for smaller samples and for data generated from a beta prior skewed towards low failure probabilities than those from a symmetric beta. However, for the large sample size, these methods showed the tendency of having the consistency property while the prior matching moments technique yielded negative bias for sample sizes bigger than 20. The estimators from the unweighted marginal matching moments yield moderate results in terms of biasedness but the same order of mean-squared error as the two most complicated methods.

The distributions of the estimated prior mean and variance were also obtained. The distribution of the prior mean estimator was found to be nearly identical for the four estimation techniques considered (excluding the prior maximum likelihood), because the mean of the beta distribution depends solely on the ratio of beta parameters b/a. However, the variance of the beta distribution with bigger values of parameters a and b has smaller variance than the distribution with the smaller parameters.

From the estimated prior distributions, the distribution of the estimated 95-th percentiles and the distribution of the fraction of the estimated prior distribution greater than the true 95-th percentile were examined. The prior matching moments method appears to be slightly more conservative from a safety viewpoint since slightly higher values of the means of both distribution are obtained with this method than with the others.

Based on the results of all the statistics considered in this study, the prior matching moments technique appeared to be the best of the five methods for estimating the beta parameters from the sample data whose size is equal to or less than 50. However, it would be interesting to investigate further for sample sizes larger than 50, because the three marginal-based techniques showed a tendency to yield comparable or better results to the prior matching moments method for large sample sizes.

6.0 REFERENCES

1. D. L. Harnett, Introduction to Statistical Methods, Addison-Wesley, 1972.

2. R. R. Fullwood, R. C. Erdmann, E. T. Rumble and G. S. Lellouche, Application of the Bayes Equation to Predicting Reactor System Reliability, Nuclear Technology, 34, Aug. 1977.

3. J. C. Kleinman, Proportions with Extraneous Variance: Single and Independent Samples, T. Am. Stat. Assoc., 68, 46, 1973.

4. A. M. Mood, F. A. Graybill and D. C. Boes, Introduction to the Theory of Statistics, McGraw-Hill, 1974.

5. Bayes formulated his work in 1763 and his paper can be found in Biometrika, 45, 293, 1958.

6. N. L. Johnson and S. Kotz, Continuous Univariate Distributions-2, Houghton Mifflin Co., 1970.

7. H. Raiffa and R. Schlaifer, Applied Statistical Design Theory, Harvard University Press, Boston, MA, 1961.

8. H. F. Martz, Jr. and R. A. Waller, The Basics of Bayesian Reliability Estimation from Attribute Test Data, Los Alamos Scientific Laboratory, Report LA-6126, Feb. 1976.

9. R. G. Krutchkoff, Empirical Bayes Estimation, The American Statistitian, 26, 14-16, December 1972.

10. J. B. Copas, Empirical Bayes Methods and the Repeated Use of a Standard, Biometrika, 59, 349-360, 1972.

11. B. S. Griffin and R. G. Krutchkoff, Optimal Linear Estimators: An Empirical Bayes Version with Application to the Binomial Distribution, Biometrika, 58, 195-201, 1969.

12. G. H. Lemon and R. G. Krutchkoff, An Empirical Bayes Smoothing Technique, Biometrika, 56, 361-365, 1969.

13. H. F. Martz, Jr. and M. G. Lian, Empirical Bayes Estimation of the Binomial Parameter, Biometrika, 61, 517-523, 1974.

14. G. H. Lemon, An Empirical Bayes Approach to Reliability, IEEE Trans. Rel. R-21, 155-158, August 1972.

References - Continued

15. J. K. Shultis and N. D. Eckhoff, Selection of Beta Prior Distribution Parameters from Component Failure Data, IEEE Transactions on Power Apparatus and Systems, Vol. PAS-98, No. 2, pp. 400-407, 1979.

16. N. L. Johnson and S. Kotz, Discrete Distributions, Houghton Mifflin Co., 1969.

17. S. L. Meyer, Data Analysis for Scientists and Engineers, Wiley, 1975.

18. R. L. Winkler and W. L. Hayes, Statistic, Probability, Inference and Decision, HRW, 1975.

19. Radiation Shielding, Analysis and Design Principles as Applied to Nuclear Defense Planning, TR 40, Nov. 1966, prepared by Office of Civil Defense and Kansas State University.

20. G. S. Fishman, Concepts and Methods in Discrete Event Digital Simulation, Wiley, 1973.

21. Scientific Subroutine Package (360-A-CM-03X), Version III Programmer's Manual, H20-0205, IBM, 1963.

22. R. A. Waller and H. F. Martz, Jr., Bayesian Reliability Estimation: The State of Art and Future Needs, Los Alamos Scientific Laboratory, Report LA-6003, Oct. 1975.

23. R. L. Winkler, Introduction to Bayesian Inference and Decision, HRW, 1972.

24. A. Hald, Statistical Theory with Engineering Application, Wiley, 1952.

25. K. S. Kolbig, Programs for Computing the Logarithm of the Gamma Function and the Digamma Function, for Complex Arguments, Comp. Phys. Comm., 4, 221, 1972.

26. A. T. de Medeiros and G. Schwachleim, Polygamma Functions with Arbitrary Precision, Comm. ACM, 12, 213, 1969.

27. O. G. Ludwig, Incomplete Beta Ratio, Comm. ACM, 6, 1963, 314; also see Collected Algorithms from CACM, Algorithm 1979 and modifications by N. E. Bosten and E. L. Battiste, 1972, and by M. C. Pike and J. Soo Hoo, 1975.

7.0. APPENDICES

APPENDIX A

Derivation of Expectation of Statistic S of Eq. (2.20)

From Eq. (2.17) a statistic S is defined by

$$S \equiv \sum_{i=1}^{N} w_i \, (\hat{p}-\hat{p}_i)^2, \qquad (A.1)$$

where

$$\hat{p}_i \equiv \frac{k_i}{n_i}.$$

From Eq. (3.19a), $\mu=E(\hat{p})$ and consequently

$$S = \sum_{i=1}^{N} w_i \, [(\hat{p}_i-\mu) - (\hat{p}-\mu)]^2,$$

or

$$S = \sum_{i=1}^{N} w_i \, [(\hat{p}_i-\mu)^2 - 2\,(\hat{p}_i-\mu)(\hat{p}-\mu) + (\hat{p}-\mu)^2].$$

Expectation of the statistic S can be computed as

$$E(S) = \sum_{i=1}^{N} w_i E(\hat{p}_i-\mu)^2 - 2 \sum_{i=1}^{N} w_i E[(\hat{p}_i-\mu)(\hat{p}-\mu)] + \sum_{i=1}^{N} w_i E(\hat{p}-\mu)^2. \qquad (A.2)$$

By definition $E(\hat{p}_i-\mu)^2$ is the variance of $\hat{p}_i$, and thus the first term of Eq. (A.2) is

$$\sum_{i=1}^{N} w_i E(\hat{p}_i-\mu)^2 \equiv \sum_{i=1}^{N} w_i \, \mathrm{var}(\hat{p}_i). \qquad (A.3)$$

Also by definition $E[(\hat{p}_i-\mu)(\hat{p}-\mu)]$ is the covariance of $\hat{p}_i$ and $\hat{p}$, and therefore the second term of Eq. (A.2) becomes

$$-2\sum_{i=1}^{N} w_i E[(\hat{p}_i-\mu)(\hat{p}-\mu)] \equiv -2\sum_{i=1}^{N} w_i \operatorname{cov}(\hat{p}_i,\hat{p})$$

$$= -2\sum_{i=1}^{N} w_i \operatorname{cov}(\hat{p}_i, \frac{w_1\hat{p}_1+w_2\hat{p}_2+ --- + w_N\hat{p}_N}{w})$$

$$= -2\sum_{i=1}^{N} w_i \operatorname{cov}(\hat{p}_i, \frac{w_i\hat{p}_i}{w})$$

$$= -2\sum_{i=1}^{N} w_i \frac{w_i}{w} \operatorname{cov}(\hat{p}_i,\hat{p}_i)$$

$$= -2\sum_{i=1}^{N} \frac{w_i^2}{w} \operatorname{var}(\hat{p}_i) \qquad (A.4)$$

Finally, the last term of Eq. (A.1) can be expressed as

$$\sum_{i=1}^{N} w_i E(\hat{p}-\mu)^2 \equiv \sum_{i=1}^{N} w_i \operatorname{var}(\hat{p})$$

$$= \operatorname{var}(\hat{p}) \sum_{i=1}^{N} w_i$$

$$= w \operatorname{var}(\hat{p})$$

$$= w \operatorname{var}(\frac{w_1\hat{p}_1+w_2\hat{p}_2+ --- + w_N\hat{p}_N}{w})$$

$$= w \sum_{i=1}^{N} \frac{w_i^2}{w^2} \operatorname{var}(\hat{p}_i)$$

$$= \sum_{i=1}^{N} \frac{w_i^2}{w} \operatorname{var}(\hat{p}_i) \qquad (A.5)$$

Upon substitution of Eqs. (A.3), (A.4) and (A.5) into Eq. (A.2), one obtains

$$E(S) = \sum_{i=1}^{N} w_i \mathrm{var}(\hat{p}_i) - 2\sum_{i=1}^{N} \frac{w_i^2}{w} \mathrm{var}(\hat{p}_i) + \sum_{i=1}^{N} \frac{w_i^2}{w} \mathrm{var}(\hat{p}_i)$$

$$= \sum_{i=1}^{N} w_i \ \mathrm{var}(\hat{p}_i) - \sum_{i=1}^{N} \frac{w_i^2}{w} \mathrm{var}(\hat{p}_i)$$

$$= \sum_{i=1}^{N} w_i (1- \frac{w_i}{w}) \ \mathrm{var}(\hat{p}_i) \ . \qquad (A.6)$$

From Eq. (2.19b), the variance of $\hat{p}_i$ is given by

$$\mathrm{var}(\hat{p}_i) = \frac{\mu(1-\mu)}{n_i} + r\mu(1-\mu) \ (1- \frac{1}{n_i}) \ . \qquad (A.7)$$

Thus, Eq. (A.6) becomes

$$E(S) = \sum_{i=1}^{N} w_i (1- \frac{w_i}{w}) \ [\frac{\mu(1-\mu)}{n_i} + r\mu(1-\mu)(1- \frac{1}{n_i})]$$

$$= \mu(1-\mu) \sum_{i=1}^{N} \frac{w_i}{n_i} (1- \frac{w_i}{w}) + r\mu(1-\mu) \sum_{i=1}^{N} (1- \frac{1}{n_i}) \ w_i (1- \frac{w_i}{w})$$

$$= \mu(1-\mu) \sum_{i=1}^{N} \frac{w_i}{n_i} (1- \frac{w_i}{w}) + r\mu(1-\mu) \ [\sum_{i=1}^{N} w_i (1- \frac{w_i}{w}) - \sum_{i=1}^{N} \frac{w_i}{n_i} (1- \frac{w_i}{w})]$$

(A.8)

Appendix B

Evaluation of Polygamma Functions

In the Newton-Raphson evaluation of the numerical solution of the maximum likelihood estimates by Eqs. (2.30a) and (2.30b), both the diagamma function and its derivative, the trigamma function, must be evaluated over a wide range of arguments. The procedure used in this study is based on a power series expansion of these functions for large arguments, and a recursion relation for small agruments [25,26].

The polygamma function $\psi^m(z)$ is defined as

$$\psi^m(z) = \frac{d^m\psi(z)}{dz^m} = \frac{d^{m+1}}{dz^{m+1}} [\ln\Gamma(z)].$$

The digamma function and trigamma functions are special cases of the polygamma function (m=0 and 1 respectively). These functions may be evaluated accurately by the formulae below:

1. Digamma (m=0):

$$z \geq 8 \qquad \psi(z) \simeq \ln z - \frac{1}{2z} - \sum_{k=1}^{10} \frac{B_{2k}}{2k} z^{-2k}$$

$$z < 8 \qquad \psi(z) = \psi(n+z) - \sum_{k=1}^{n} (z+k-1)^{-1}$$

where B_{2k} are the Bernoulli numbers.

2. Trigamma (m=1):

$$z \geq 8 \qquad \psi^1(z) \simeq \frac{1}{2} + \frac{1}{2z^2} + \sum_{k=1}^{10} B_{2k} z^{-(2k+1)}$$

$$z < 8 \qquad \psi^1(z) = \psi^1(n+z) + \sum_{k=1}^{n} (z+k-1)^{-2}$$

3. Polygamma ($m>1$):

$$z \geq 8 \qquad \psi^{m}(z) = (-1)^{m-1}\left[\frac{(m-1)!}{z^{m}} + \frac{m!}{2z^{m+1}} + \sum_{k=1}^{10} B_{2k}\,\frac{(2k+m-1)!}{(2k)!}\,z^{-(2k+m)}\right]$$

$$z < 8 \qquad \psi^{m}(z) = \psi^{m}(2+n) - (-1)^{m}\, m! \sum_{k=1}^{n} (z+k-1)^{-m-1}$$

APPENDIX C

Evaluation of the Incomplete Beta Functions

The incomplete beta function $I_p(x,y)$ is calculated from the following expression [27]:

$$I_p(x,y) = \frac{\text{INFSUM}\ p^x\, \Gamma(PS+x)}{\Gamma(PS)\ \Gamma(x+1)} + \frac{P^x\ (1-P)^y\ \Gamma(x+y)\ \text{FINSUM}}{\Gamma(x)\ \Gamma(y+1)}$$

where INFSUM and FINSUM represent two series summations defined as follows:

$$\text{INFSUM} = \sum_{j=1}^{\infty} \frac{x(1-PS)_j}{x+j} \frac{p^j}{j!} \text{ , where}$$

$$(1-PS) = \begin{cases} 1 \text{ , } j = 0 \\ \Gamma(1+y-PS)/\Gamma(1-PS) \text{ , } j > 0 \end{cases}$$

and

$$\text{FINSUM} = \sum_{j=1}^{[y]} \frac{y(y-1)\ldots(y-j+1)}{(x+y-1)(x+y-2)\ldots(x+y-j)} \frac{1}{(1-p)^j}$$

where [y] is equal to the largest integer less than y. If [y]=0, the FINSUM=0. The quantity PS is defined as

$$PS = \begin{cases} 1 \quad \text{if } y \text{ is integer} \\ y - [y]\text{, otherwise.} \end{cases}$$

The above algorithm (combined with scaling to avoid numerical inaccuracies encountered when using the gamma function with large arguments) was incorporated into a FORTRAN program MDBETA by Bosten and Battiste [27]. This program (modified in accordance to remarks made by Pike and Soo Hoo [27]) was used in the present analysis.

The program MDBETA is significantly more accurate than the widely used program BDTR [21], especially at large arguments. For example, in the case p=0.5, x=y=2000, MDBETA gives the correct value, 0.5, while BDTR gives 0.497026.

APPENDIX D

Computer Program Listings

Listing of Computer Program Used to Generate Simulated Failure Data

(a=1.2, b=23)

```
C*** BUSIM-RANDU ******************************************************
C*  PURPOSE : GENERATE RANDOM VARIATE FROM DISCRETE UNIFORM           *
C*            DIST. RANGING FROM L1 TO L2                             *
C*            AND BETA-BINOMIAL DISTRIBUTION WITH PARAMETERS A & B    *
C*            BY USING SUBROUTINE RANDU                               *
C*  INPUT :                                                           *
C*   CARD 1 (I10)                                                     *
C*    IX      STARTING VALUE FOR SUBROUTINE RANDU                     *
C*            (SEE SUBROUTINE RANDU FOR DETAILS)                      *
C*   CARD 2 (4I5.2G10.3)                                              *
C*    L1.L2   LOWER & UPPER LIMITS OF THE DISCRETE UNIFORM DIST.      *
C*    NS      NO. OF SAMPLES REQUIRED                                 *
C*    NSS     SAMPLE SIZE                                             *
C*  SUBROUTIVE REQUIRED : RANDU                                       *
C*  NOTE : THIS ROUTINE IS USED FOR PARAMETERS A=1.2.B=23.0           *
C*********************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION NT(30).KF(30)
      PRINT 600
600   FORMAT('1')
C***READ STARTING VALUE FOR RANDU
      READ 501.IX
501   FORMAT(I10)
      PRINT 601.IX
501   FORMAT(T5,'RANDU STARTING VALUE = ',I9/)
100   READ(5.500.END=999)L1.L2.NS.NSS.A.B
C***NS   NO.OF SAMPLE REQUIRED
C***NSS  SAMPLE SIZE
500   FORMAT(4I5.2G10.3)
      PRINT 605.A.B.L1.L2
605   FORMAT(T5,'SIMULATION'/
     *T5,'FROM BETA-BINOMIAL DISTRIBUTION'/
     *T5,'WITH PARAMETERS : A = ',G14.7/
     *T5,'                  B = ',G14.7/
     *T5,'AND DISCRETE UNIFORM DISTRIBUTION'/
     *T5,'RANGING FROM',1X,I3,2X,'TO',2X,I3/)
      P =1.0D00/(L2-L1+1)
      DO 900 INS=1.NS
      DO 700 INSS=1.NSS
C***GENERATE RANDOM NUMBER
      CALL RANDU(IX.IY.U)
      IX=IY
C***TRANSFORM TO DISC.UNI.DIST.
      UP=U/P
      IP=UP
      UIP=IP
      NT(INSS)=L1+IP
      IF(UIP.EQ.UP.AND.U.NE.0.0D00) NT(INSS)=NT(INSS)-1
C***GENERATE RANDOM NUMBER
      CALL RANDU(IX.IY.U)
      IX=IY
C***TRANSFORM TO BETA-BINOMIAL DIST.(METHOD II)
      KF(INSS)=0
      AG1=NT(INSS)+B
      AG2=NT(INSS)+A+B
      BB1=DLGAMA(A+B)+DLGAMA(AG1)-DLGAMA(B)-DLGAMA(AG2)
      BB=DEXP(BB1)
      CB=BB
```

```
200    IF(U.LE.CB) GO TO 700
       AG1=(KF(INSS)+A)*(NT(INSS)-KF(INSS))
       BB=BB*AG1/((NT(INSS)-KF(INSS)+B-1)*(KF(INSS)+1))
       CB=CB+BB
       KF(INSS)=KF(INSS)+1
       GO TO 200
700    CONTINUE
C***PUNCH FAILURE ATTRIBUTE DATA
C***NT   NO.OF TRIES
C***KF   NO.OF FAILURES
       PUNCH 801,(NT(I),KF(I),I=1,NSS)
801    FORMAT(20I4)
900    CONTINUE
       GO TO 100
999    PRINT 600
       STOP
       END
```

```
C
C     ...............................................................
C
C        SUBROUTINE RANDU
C
C        PURPOSE
C           COMPUTES UNIFORMLY DISTRIBUTED RANDOM REAL NUMBERS BETWEEN
C           0 AND 1.0 AND RANDOM INTEGERS BETWEEN ZERO AND
C           2**31. EACH ENTRY USES AS INPUT AN INTEGER RANDOM NUMBER
C           AND PRODUCES A NEW INTEGER AND REAL RANDOM NUMBER.
C
C        USAGE
C           CALL RANDU(IX,IY,YFL)
C
C        DESCRIPTION OF PARAMETERS
C           IX - FOR THE FIRST ENTRY THIS MUST CONTAIN ANY ODD INTEGER
C                NUMBER WITH NINE OR LESS DIGITS. AFTER THE FIRST ENTRY,
C                IX SHOULD BE THE PREVIOUS VALUE OF IY COMPUTED BY THIS
C                SUBROUTINE.
C           IY - A RESULTANT INTEGER RANDOM NUMBER REQUIRED FOR THE NEXT
C                ENTRY TO THIS SUBROUTINE. THE RANGE OF THIS NUMBER IS
C                BETWEEN ZERO AND 2**31
C           YFL- THE RESULTANT UNIFORMLY DISTRIBUTED, FLOATING POINT,
C                RANDOM NUMBER IN THE RANGE 0 TO 1.0
C
C        REMARKS
C           THIS SUBROUTINE IS SPECIFIC TO SYSTEM/360 AND WILL PRODUCE
C           2**29 TERMS BEFORE REPEATING.  THE REFERENCE BELOW DISCUSSES
C           SEEDS (65539 HERE), RUN PROBLEMS, AND PROBLEMS CONCERNING
C           RANDOM DIGITS USING THIS GENERATION SCHEME.  MACLAREN AND
C           MARSAGLIA, JACM 12, P. 83-89, DISCUSS CONGRUENTIAL
C           GENERATION METHODS AND TESTS.  THE USE OF TWO GENERATORS OF
C           THE RANDU TYPE, ONE FILLING A TABLE AND ONE PICKING FROM THE
C           TABLE, IS OF BENEFIT IN SOME CASES.  65549 HAS BEEN
C           SUGGESTED AS A SEED WHICH HAS BETTER STATISTICAL PROPERTIES
C           FOR HIGH ORDER BITS OF THE GENERATED DEVIATE.
C           SEEDS SHOULD BE CHOSEN IN ACCORCANCE WITH THE DISCUSSION
C           GIVEN IN THE REFERENCE BELOW.  ALSC, IT SHOULD BE NOTED THAT
C           IF FLOATING POINT RANDOM NUMBERS APE DESIRED,AS ARE
C           AVAILABLE FROM RANDU, THE RANDOM CHARACTERISTICS OF THE
C           FLOATING POINT DEVIATES ARE MODIFIED AND IN FACT THESE
C           DEVIATES HAVE HIGH PROBABILITY OF HAVING A TRAILING LOW
C           ORDER ZERO BIT IN THEIR FRACTICNAL PART.
C
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
C           NONE
C
C        METHOD
C           POWER RESIDUE METHOD DISCUSSED IN IBM MANUAL C20-8011,
C           RANDOM NUMBER GENERATION AND TESTING
C
C     ...............................................................
C
      SUBROUTINE RANDU(IX,IY,YFL)
      REAL*8 YFL
      IY=IX*65539
      IF(IY)5,6,6
    5 IY=IY+2147483647+1
    6 YFL=IY
      YFL=YFL*.4656613E-9
      RETURN
      END
```

```
C*** SUPPLY-BUSIM-RANDU 2 ***********************************************
C*  PURPOSE : COMPUTE MEAN,PROBABILITY FUNCTION,CUM.PROB. FUNCTION     *
C*            OF THE BETA-BINOMIAL DIST.                               *
C*            TO USE WITH THE ROUTINE BUSIM-RANDU 2                    *
C*  INPUT :                                                            *
C*   CARD 1 (2G10.5,2I5)                                               *
C*    A.B     PARAMETERS OF THE BETA-BINOMIAL DIST.                    *
C*    NI,NJ   LOWER & UPPER LIMITS OF THE DISCRETE UNIFORM DIST.       *
C***********************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION C(300).F(300),KF(300),NF(300)
      READ 500,A.B.NI,NJ
500   FORMAT(2G10.5,2I5)
      NN=NJ-NI+1
      DO 100 I=1.NN
      N=NI+I-1
      NF(I)=N
      AVG=N*A/(A+B)
      KAVG=AVG
      K=0
      AG1=N+B
      AG2=N+A+B
      BB1=DLGAMA(A+B)+DLGAMA(AG1)-DLGAMA(B)-DLGAMA(AG2)
      BB=DEXP(BB1)
      CB=BB
200   IF(K.GE.KAVG) GO TO 210
      AG1=(K+A)*(N-K)/((N-K+B-1)*(K+1))
      BB=BB*AG1
      CB=CB+BB
      K=K+1
      GO TO 200
210   C(I)=CB
      F(I)=BB
      KF(I)=K
100   CONTINUE
C***PUNCH OUTPUT
C***KF    MEAN OF THE BETA-BINOMIAL DIST.
C***NF    NO.OF TRIES
C***F     PROB.FUNCTION
C***C     CUM.PROB. FUNCTION
      PUNCH 800,(KF(I).NF(I).F(I),C(I).I=1.NN)
800   FORMAT(2I5,2D25.16)
      STOP
      END
```

Listing of Computer Program Used to Generate Simulated Failure Data

(a=b=5)

```
C*** BUSIM-RANDU 2 ***********************************************************
C*  PURPOSE : GENERATE RANDOM VARIATE FROM DISCRETE UNIFORM                 *
C*            DIST. RANGING FROM L1 TO L2                                   *
C*            AND BETA-BINOMIAL DISTRIBUTION WITH PARAMETERS A & B          *
C*            BY USING SUBROUTINE RANDU                                     *
C*  INPUT :                                                                 *
C*   CARD 1 (I10)                                                           *
C*    IX      STARTING VALUE FOR SUBROUTINE RANDU                           *
C*            (SEE SUBROUTINE RANDU FOR DETAILS)                            *
C*   CARD 2 (4I5,2G10.3)                                                    *
C*    L1,L2   LOWER & UPPER LIMITS OF THE DISCRETE UNIFORM DIST.            *
C*    NS      NO. OF SAMPLES REQUIRED                                       *
C*    NSS     SAMPLE SIZE                                                   *
C*   CARD 3 (2I5,2D25.16)                                                   *
C*    K,N,F,C OUTPUT FROM THE ROUTINE SUPPLY-BUSIM-RANDU 2                  *
C*  SUBROUTINE REQUIRED : RANDU                                             *
C*  NOTE : THIS ROUTINE IS USED FOR PARAMETERS A=B= 5.0                     *
C***************************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION NT(30),KF(30)
      DIMENSION C(300),K(300),F(300),N(300)
      PRINT 600
600   FORMAT('1')
C***READ STARTING VALUE FOR RANDU
      READ 501,IX
501   FORMAT(I10)
      PRINT 601,IX
601   FORMAT(T5,'RANDU STARTING VALUE = ',I9/)
100   READ(5,500,END=999)L1,L2,NS,NSS,A,B
C***NS   NO.OF SAMPLE REQUIRED
C***NSS  SAMPLE SIZE
500   FORMAT(4I5,2G10.3)
      NN=L2-L1+1
      READ 510,(K(I),N(I),F(I),C(I),I=1,NN)
510   FORMAT(2I5,2D25.16)
      PRINT 605,A,B,L1,L2
605   FORMAT(T5,'SIMULATION'/
     *T5,'FROM BETA-BINOMIAL DISTRIBUTION'/
     *T5,'WITH PARAMETERS : A = ',G14.7/
     *T5,'                  B = ',G14.7/
     *T5,'AND DISCRETE UNIFORM DISTRIBUTION'/
     *T5,'RANGING FROM',1X,I3,2X,'TO',2X,I3/)
      P =1.0D00/(L2-L1+1)
      DO 900 INS=1,NS
      DO 700 INSS=1,NSS
C***GENERATE RANDOM NUMBER
      CALL RANDU(IX,IY,U)
      IX=IY
C***TRANSFORM TO DISC.UNI.DIST.
      UP=U/P
      IP=UP
      UIP=IP
      NT(INSS)=L1+IP
      IF(UIP.EQ.UP.AND.U.NE.0.0D00) NT(INSS)=NT(INSS)-1
C***GENERATE RANDOM NUMBER
      CALL RANDU(IX,IY,U)
      IX=IY
C***TRANSFORM TO BETA-BINOMIAL DIST.(METHOD III)
```

```
      NX=NT(INSS)-L1+1
      KF(INSS)=K(NX)
      IF(NT(INSS).EQ.N(NX)) GO TO 110
      PRINT 610,N(NX)
610   FORMAT(I2,'*** DATA MISORDER AT N = ',I4,' ***')
      STOP
110   BB=F(NX)
      CB=C(NX)
      IF(U-C(NX)) 115,700,125
115   CB=CB-BB
116   IF(U.GT.CB) GO TO 700
      AG1=(NT(INSS)-KF(INSS)+B)*KF(INSS)
      BB=BB*AG1/((KF(INSS)-1+A)*(NT(INSS)-KF(INSS)+1))
      CB=CB-BB
      KF(INSS)=KF(INSS)-1
      GO TO 116
125   AG1=(KF(INSS)+A)*(NT(INSS)-KF(INSS))
      BB=BB*AG1/((NT(INSS)-KF(INSS)-1+B)*(KF(INSS)+1))
      CB=CB+BB
      KF(INSS)=KF(INSS)+1
      IF(U.LE.CB) GO TO 700
      GO TO 125
700   CONTINUE
C***PUNCH FAILURE ATTRIBUTE DATA
C***NT   NO.OF TRIES
C***KF   NO.OF FAILURES
      PUNCH 801,(NT(I),KF(I),I=1,NSS)
801   FORMAT(20I4)
900   CONTINUE
      GO TO 100
999   PRINT 600
      STOP
      END
```

Listing of Computer Program Used to Compute Beta Parameter Estimates

```
C********************************* BETA III ************************************
C*                                                                             *
C* THIS PROGRAM                                                                *
C*   - CALCULATES THE PARAMETERS A AND B OF AN ASSUMED BETA MIXING             *
C* DISTRIBUTION BY  FOUR TECHNIQUES:  (1) MATCHING MOMENTS OF THE EXPERIMENTAL *
C* DATA TO THOSE OF THE MARGINAL DISTRIBUTION,    (2) MATCHING MOMENTS OF THE  *
C* DATA TO THOSE OF THE PRIOR DISTRIBUTION,        (3) THE MAXIMUM LIKELIHOOD  *
C* METHOD WITH BETA-BINOMIAL DISTRIBUTION,AND (4) THE MAXIMUM LIKELIHOOD       *
C* METHOD WITH BETA DISTRIBUTION                                               *
C*   - ALSO CALCULATES AND PLOTS BETA DISTRIBUTION (BOTH PROBABILITY DENSITY   *
C* FUNCTION AND CUMULATIVE DISTRIBUTION FUNCTION)                              *
C* FOR EACH METHOD AND COMPARISON OF FOUR METHODS.                             *
C*                                                                             *
C* INPUT DATA:                                                                 *
C*                                                                             *
C*    CARD 1    (20A4)                                                         *
C*       TITLE  = THE TITLE OF THE PROBLEM (80 COLUMNS)                        *
C*                                                                             *
C*    CARD 2    (3I5,5G10.3)                                                   *
C*       NITER  = MAXIMUM NUMBER OF ITERATIONS FOR METHOD 1 AND FOR NUMERICAL  *
C*                SOLUTION IN METHOD 3&4.IF =0 ONLY MOMENTS METHODS            *
C*                CALCULATIONS ARE PERFORMED.                                  *
C*       IOUT   = 1 IF INTERMEDIATE OUTPUT IS DESIRED FOR THE ITERATIONS IN    *
C*                METHOD 1 AND FOR THE NUMERICAL SOLUTION IN METHOD 3&4;IF =0  *
C*                ONLY FINAL RESULTS FOR ALL FOUR  METHODS ARE PRINTED OUT.    *
C*       IPROB  = 1 IF A COMPARISON OF THE CLASSICAL AND BAYESIAN FAILURE      *
C*                PROBABILITIES FOR EACH COMPONENT IS DESIRED;  IF =0 THIS     *
C*                OPTION IS BYPASSED.                                          *
C*       Y1     = INITIAL GUESS FOR A IN METHOD 3;  IF =0 RESULT FROM METHOD 2 *
C*                WILL BE USED FOR INITIAL GUESS.                              *
C*       Y2     = INITIAL GUESS FOR B IN METHOD 3;  IF =0 RESULT FROM METHOD 2 *
C*                WILL BE USED FOR INITIAL GUESS.                              *
C*       EPS    = CONVERGENCE PARAMETER FOR METHODS 1,3 & 4.    IN METHOD 1    *
C*                ITERATIONS CONTINUE UNTIL PRIOR MEAN CHANGES BY LESS THAN    *
C*                EPS.IN METHOD 3&4 NEWTON-RAPHSON ITERATIONS CONTINUE UNTIL   *
C*                DERIVATIVES ARE < EPS.                                       *
C*       Z1     = INITIAL GUESS FOR A IN METHOD 4;  IF =0 RESULT FROM METHOD 2 *
C*                WILL BE USED FOR INITIAL GUESS.                              *
C*       Z2     = INITIAL GUESS FOR B IN METHOD 4;  IF =0 RESULT FROM METHOD 2 *
C*                WILL BE USED FOR INITIAL GUESS.                              *
C*                                                                             *
C*    CARD 3    (4G10.3,7I5)                                                   *
C*       PI,PJ,PK,PF1,NI,NJ,NL,IXOUT,IVAL,IPL,IBETA                            *
C*       IBETA  = 0; COMPUTED VALUES & PLOTS OF BETA DISTRIBUTIONS ARE         *
C*                   IGNORED.                                                  *
C*       IBETA  = 1; COMPUTED VALUES & PLOTS OF BETA DISTRIBUTIONS ARE         *
C*                   DISPLAYED ( SEE IVAL & IPL FOR MORE DETAILS).             *
C*       SEE MORE EXPLANATION IN SUBROUTINE BETDIS.                            *
C*                                                                             *
C*    CARD 4...  (16I5)                                                        *
C*       NN     = NUMBER OF PAIRS OF DATA POINTS TO BE READ                    *
C*       N(I),K(I) = NUMBER OF TRIES, NUMBER OF FAILURES FOR I-TH PLANT        *
C*                   NN PAIRS OF N(I) AND K(I) ARE TO BE ENTERED.              *
C*                                                                             *
C* SUBROUTINES REQUIRED:                                                       *
C*    NEWRAL - NEWTON-RAPHSON SOLUTION OF TWO SIMULTANEOUS EQUATIONS           *
C*    FNDATA - READS IN STARTS AND FAILURES, N(I) AND K(I). ALSO CALCULATES    *
C*              THE LIKELIHOOD FUNCTION AND ITS DERIVATIVES                    *
```

```
C*              (BETA-BINOMIAL DISTRIBUTION)                                     *
C*    FBT    - CALCULATES THE LIKELIHOOD FUNCTICN AND ITS DERIVATIVES           *
C*              (BETA DISTRIBUTIIN)                                              *
C*    POLGAM - CALCULATES THE POLYGAMMA FUNCTION                                 *
C*    VARMLE - CALCULATES VARIANCES AND COVARIANCE OF MAXIMUM LIKELIHOOD        *
C*             ESTIMATORS (EXACT EXPECTATION VALUES;BETA-BINOMIAL DIST.)        *
C*    APPMLE - CALCULATES VARIANCES AND COVARIANCE OF MAXIMUM LIKELIHOOD        *
C*             ESTIMATORS (APPROX. EXPECTATION VALUES;BETA-BINCMIAL DIST.)      *
C*    BETDIS - CALCULATE AND PLOT BETA DISTRIBUTION (PROBABILITY DENSITY        *
C*             FUNCTION AND CUMULATIVE DISTRIBUTION FUNCTICN)                   *
C*    GPA    - USED IN SUBROUTINE BETDIS                                        *
C*    PLOT   - USED IN SUBROUTINE BETDIS                                        *
C*    MDBETA - USED IN SUBROUTINE BETDIS                                        *
C*                                                                              *
C* REMARKS:                                                                     *
C*    DIMENSION OF P,PB,W,N,K ARE NN                                            *
C*                                                                              *
C********************************************************************************
      REAL*8 Y1,Y2,AA,BB,EPS,F,G,MEAN,SIG,P,PB(50),DFLOAT
      REAL*8 SIGA,SIGB,DSQRT,VARP,VARSIG,A(4),TITLE(20),DABS
      REAL*8 RBAR,W(50) ,WW,PBAR,S,QBAR,SUM1,SUM2,SSS,BA,PPBAR
      REAL*8 HEMT1(20),HEMT2(20),HEMT3(20),HEADT(4,20),DA(4),DB(4)
      REAL*8 PI,PJ,PK,V11,V22,V12,W11,W22,W12,PF1
      REAL*8 Z1,Z2,HEMT4(20)
      REAL*8 VARA,VARB,VARAND,VARBND,SIGAND,SIGBND
      REAL*8 XPBAR,XQBAR,XS,XPQ,XSUM,XSIG,XRBAR,XAA,XBB
      COMMON/DATA/NN,N(50) ,K(50)
      COMMON /Z/ P(50)
      EXTERNAL FNDER,FBT
      DATA HEMT1/'MATC','HING',' DAT','A MO','MENT','S TO',' MAR','GINA'
     *,'L DI','STRI','BUTI','ON M','CMEN','TS  ',6*'    '/
      DATA HEMT2/'MATC','HING',' DAT','A MO','MENT','S TO',' PRI','OR D'
     *,'ISTR','IBUT','ION ','MOME','NTS ',7*'    '/
      DATA HEMT3/'MAXI','MUM ','LIKE','LIHO','OD M','ETHO','D WI','TH B'
     *,'ETA-','BINO','MIAL',' DIS','T.  ',7*'    '/
      DATA HEMT4/'MAXI','MUM ','LIKE','LIHO','OD M','ETHO','D WI','TH B'
     *,'ETA ','DIST','.   ',9*'    '/
C
C**** READ IN THE PROBLEM TITLE AND DATA
   99 READ (5,12,END=98) (TITLE(I),I=1,20)
   12 FORMAT(20A4)
      PRINT 13,(TITLE(I),I=1,20)
   13 FORMAT('1',20A4)
      READ 10,NITER,IOUT,IPROB,Y1,Y2,EPS,Z1,Z2
   10 FORMAT(3I5,5G10.3)
      READ 150,PI,PJ,PK,PF1,NI,NJ,NL,IXOUT,IVAL,IPL,IBETA
150   FORMAT(4G10.3,7I5)
      CALL FNDATA(Y1,Y2,F,G,A)
      PRINT 14, (N(I),I=1,NN)
      PRINT 17, (K(I),I=1,NN)
   14 FORMAT(5X,'TRIES:     ',23I5,/(15X,23I5))
   17 FORMAT(5X,'FAILURES: ',23I5,/(15X,23I5))
      NITE=NITER
      NOM=0
C
C*** CALCULATE THE PRIOR PARAMETERS BY MATCHING DATA MOMENTS TO MARGINAL DISTRI-
C*** BUTIONS MOMENTS.
      PRINT 610
```

```
610   FORMAT('0MATCHING MOMENTS OF DATA TO THOSE OF MARGINAL DISTRIBUTIO
     *N:')
      NMAX=NITER
      IF(NMAX.EQ.0) NMAX=20
      ITER=0
      MCONV=0
      DO 51 I=1,NN
   51 P(I)=DFLOAT(K(I))/N(I)
      XPBAR=0.0D0
      DO 300 I=1,NN
300   XPBAR=XPBAR+P(I)
      XPBAR=XPBAR/NN
      XQBAR=1.0D0-XPBAR
      XS=0.0D0
      DO 305 I=1,NN
305   XS=XS+(P(I)-XPBAR)**2
      XPQ=XPBAR*XQBAR*(NN-1)
      XSUM=0.0D0
      DO 310 I=1,NN
310   XSUM=XSUM+1.0D0/N(I)
      XSUM=XSUM*XPQ/NN
      XRBAR=(XS-XSUM)/(XPQ-XSUM)
      IF(XRBAR.LE.0.0D0) GO TO 315
      XSUM=XRBAR*XPBAR*XQBAR
      XSIG=DSQRT(XSUM)
      IF(XPBAR*XQBAR.LE.XSUM) GO TO 316
      XPQ=1.0D0/XRBAR-1.0D0
      XAA=XPBAR*XPQ
      XBB=XQBAR*XPQ
      PRINT 612,XPBAR,XSIG,XAA,XBB
612   FORMAT(' NO WEIGHTING        :  MEAN=',G15.8,'     SIGMA=',G15.8,
     *';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
      GO TO 320
315   PRINT 613,XPBAR
613   FORMAT(' NO WEIGHTING        :  MEAN=',G15.8,'     R IS NEGATIVE')
      GO TO 320
316   PRINT 614,XPBAR,XSIG
614   FORMAT(' NO WEIGHTING        :  MEAN=',G15.8,'     SIGMA=',G15.8,
     *';',7X,'PRIOR PARAMETERS: A&B ARE NEGATIVE')
320   CONTINUE
      PPBAR=10.0D0
      RBAR=0.0D0
   50 ITER=ITER+1
C*** CALCULATE THE WEIGHTS
      WW=0.0D0
      DO 52 I=1,NN
      W(I)=N(I)/(1.0D0+RBAR*(N(I)-1))
   52 WW=WW+W(I)
C*** CALCULATE PBAR AND S
      PBAR=0.0D0
      DO 53 I=1,NN
   53 PBAR=PBAR+W(I)*P(I)
      PBAR=PBAR/WW
      QBAR=1.0D0-PBAR
      S=0.0D0
      DO 54 I=1,NN
   54 S=S + W(I)*(P(I)-PBAR)**2
      S=(NN-1)*S/NN
```

```
C*** CALCULATE MEAN OF PRIOR AND RBAR
      SUM1=0.0D0
      SUM2=0.0D0
      DO 55 I=1,NN
      SSS=W(I)*(1.0D0-W(I)/WW)
      SUM1=SUM1+SSS/N(I)
   55 SUM2=SUM2+SSS
      RBAR=(S-PBAR*QBAR*SUM1)/(PBAR*QBAR*(SUM2-SUM1))
      IF (RBAR.LE.0.0D0) RBAR=0.0D0
C*** CHECK FOR CONVERGENCE
      SSS=DABS((PBAR-PPBAR)/PBAR)
      IF(SSS.LE.EPS) MCONV=1
      PPBAR=PBAR
C*** CALCULATE THE A AND B PARAMETERS OF THE PRIOR DISTRIBUTION
      IF(RBAR) 56,56,57
   57 AA=PBAR*(1.0D0/RBAR-1.0D0)
      BB=QBAR*(1.0D0/RBAR-1.0D0)
      SIG=DSQRT(RBAR*PBAR*QBAR)
      IF (ITER.GT.2) GO TO 59
      IF (ITER.EQ.1) PRINT 65,PBAR,SIG,AA,BB
      IF (ITER.EQ.2) PRINT 66,PBAR,SIG,AA,BB
      GO TO 81
   59 IF(IOUT.EQ.1) PRINT 69,PBAR,SIG,AA,BB
   81 IF (MCONV.EQ.1) PRINT 67,PBAR,SIG,AA,BB
      IF((ITER.EQ.NMAX).AND.(MCONV.EQ.0)) PRINT 68,PBAR,SIG,AA,BB
      IF((MCONV.EQ.1).OR.(ITER.EQ.NMAX)) GO TO 85
      GO TO 50
   56 BA=1.0D0/PBAR - 1.0D0
      IF(ITER.GT.2) GO TO 61
      IF (ITER.EQ.1) PRINT 75,PBAR,BA
      IF (ITER.EQ.2) PRINT 76,PBAR,BA
      GO TO 82
   61 IF(IOUT.EQ.1) PRINT 79,PBAR,BA
   82 IF (MCONV.EQ.1) PRINT 77,PBAR,BA
      IF((ITER.EQ.NMAX).AND.(MCONV.EQ.0)) PRINT 78,PBAR,BA
   65 FORMAT(' BINOMIAL WEIGHTING :  MEAN=',G15.8,'      SIGMA=',
     2G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
   66 FORMAT(' EMPIRICAL WEIGHTING:  MEAN=',G15.8,'      SIGMA=',
     1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
   67 FORMAT(' CONVERGED RESULT    :  MEAN=',G15.8,'      SIGMA=',
     1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
   68 FORMAT(' NO CONVERGENCE      :  MEAN=',G15.8,'      SIGMA=',
     1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
   69 FORMAT(23X,'MEAN=',G15.8,'      SIGMA=',
     1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
   75 FORMAT(' BINOMIAL WEIGHTING :  MEAN=',G15.8,'      SIGMA=',
     23X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
   76 FORMAT(' EMPIRICAL WEIGHTING:  MEAN=',G15.8,'      SIGMA=',
     13X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
   77 FORMAT(' CONVERGED RESULT    :  MEAN=',G15.8,'      SIGMA=',
     13X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
   78 FORMAT(' NO CONVERGENCE      :  MEAN=',G15.8,'      SIGMA=',
     13X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
   79 FORMAT(20X,'   MEAN=',G15.8,'      SIGMA=',
     13X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
85    IF(MCONV.NE.1.OR.RBAR.LE.0.0D00) GO TO 86
      NUM=NUM+1
      DA(NUM)=AA
      DB(NOM)=BB
      DO 110 I=1,20
110   HEADT(NOM,I)=HEMT1(I)
C
C**** CALCULATE A AND B BY MATCHING THE DATA MOMENTS TO THOSE OF THE PRIOR.
86    MEAN=0.0D00
      SIG4=0.0D00
      SIG=0.0D0
      DO 34 I=1,NN
      P(I)=DFLOAT(K(I))/N(I)
   34 MEAN=MEAN+P(I)
      MEAN= MEAN/NN
      DO 35 I=1,NN
      SIG4=SIG4+(P(I)-MEAN)**4
   35 SIG=SIG + (P(I)-MEAN)**2
      SIG=SIG/(NN-1)
      SIG4=SIG4/NN
      AA=(MEAN**2/SIG)*(1.0D0-MEAN) - MEAN
      BB=(1.0D0-MEAN)*AA/MEAN
      SSS=DSQRT(SIG)
      IF(NN.LE.2) GO TO 40
      VARP=SIG/NN
      VARSIG=2.0D0*SIG**2/(NN-1)
      VAR=(SIG4-(NN-3)*SIG**2/(NN-1))/NN
      VARA=       ((((2.0D0-3.0D0*MEAN)*MEAN/SIG)-1.0D0)**2)*VARP
     1  + VARSIG*(MEAN**2*(1.0D0-MEAN)/SIG**2)**2
      SIGA=DSQRT(VARA)
      VARB=VARP*((((1.0D0-4.0D0*MEAN+3.0D0*MEAN**2)/SIG)+1.0D0)**2 +
     2 VARSIG*((MEAN*(1.0D0-MEAN)**2)/(SIG**2))**2
      SIGB=DSQRT(VARB)
      VARAND=       ((((2.0D00-3.0D00*MEAN)*MEAN/SIG)-1.0D00)**2)*VARP
     3  +VAR*(MEAN**2*(1.0D00-MEAN)/SIG**2)**2
      SIGAND=DSQRT(VARAND)
      VARBND=VARP*((((1.0D0-4.0D0*MEAN+3.0D0*MEAN**2)/SIG)+1.0D0)**2 +
     2    VAR*((MEAN*(1.0D0-MEAN)**2)/(SIG**2))**2
      SIGBND=DSQRT(VARBND)
      PRINT 37,MEAN,SSS,AA,BB
      PRINT 38,VARA,VARB,SIGA,SIGB,VARAND,VARBND,SIGAND,SIGBND
38    FORMAT(' VARIANCE AND STANDARD DEVIATION ESTIMATES (ASSUMING NORMA
     *L DISTRIBUTION) :',T92,'VAR(A)=',G13.6,'VAR(B)=',G13.6/
     *T92,'SIG(A)=',G13.6,'SIG(B)=',G13.6/
     *' VARIANCE AND STANDARD DEVIATION ESTIMATES (DISTRIBUTION INDEPEND
     *ENT) :',T92,'VAR(A)=',G13.6,'VAR(B)=',G13.6/
     *T92,'SIG(A)=',G13.6,'SIG(B)=',G13.6)
      GO TO 39
   40 PRINT 37, MEAN,SSS,AA,BB
   37 FORMAT('0',//'0MATCHING MOMENTS OF THE DATA TO THOSE OF THE PRIOR
     1DISTRIBUTION:',/' PRIOR MOMENTS:',8X,'MEAN=',G15.8,5X,'SIGMA=',
     2G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'    B=',G15.8)
39    NOM=NOM+1
      DA(NOM)=AA
      DB(NOM)=BB
      DO 120 I=1,20
120   HEADT(NOM,I)=HEMT2(I)
C
C*** CALCULATE A AND B BY MAX.LIKELIHOOD METHOD WITH BETA-BINOMIAL DISTRIBUTION
      IF(NITER.EQ.0) GO TO 41
      IF(Y1.EQ.0.0D0) GO TO 32
      PRINT 11, Y1,Y2,EPS,NITER
   11 FORMAT('0',/'OMAXIMUM LIKELIHCCD METHCD WITH BETA-BINOMIAL DISTRIB
     *UTION:',
     1/5X,'INITIAL STARTING POINTS',2G15.8,/5X,'ACCURACY PARAMETER=',
     2G12.4,/5X,'MAXIMUM NUMBER OF ITERATIONS=',I4)
      GO TO 33
   32 Y1=AA
      Y2=BB
      PRINT 36,Y1,Y2,EPS,NITER
   36 FORMAT('0',/'OMAXIMUM LIKELIHOOD METHGD WITH BETA-BINGMIAL DISTRIB
     *UTION:',
     1/5X,'INITIAL STARTING POINTS CALCULATED BY MATCHING MCMENTS TO PRI
     2OR',2G15.8,/5X,'ACCURACY PARAMETER=',G12.4,/5X,'MAXIMUM NUMBER OF
     3ITERATIONS=',I4)
C*   SOLVE FOR A AND B BY THE NEWTON-RAPHSON METHOD
   33 IOT=IOUT
      CALL NEWRAL(Y1,Y2,F,G,FNDER,EPS,NITER,IOT)
      MEAN=Y1/(Y1+Y2)
      SIG=DSQRT(Y1*Y2/((Y1+Y2+1)*(Y1+Y2)**2))
      IF (IOT)  15,20,15
   15 PRINT 16,Y1,Y2,IOT
   16 FORMAT(5X,'SOLUTION CONVERGED TO:   A=',G15.8,'   AND   B=',G15.8,
     1'    AFTER',I3,' ITERATIONS.')
      PRINT 24, MEAN,SIG,Y1,Y2
   24 FORMAT(' PRIOR MOMENTS:',8X,'MEAN=',G15.8,'     SIGMA=',
     1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
C *** CALCULATE VARIANCES & COVARIANCE OF MAX. LIKELIHOOD ESTIMATORS
      CALL VARMLE(Y1,Y2,NN,N,V11,V22,V12)
      CALL APPMLE(Y1,Y2,NN,N,K,W11,W22,W12)
      NOM=NOM+1
      DA(NOM)=Y1
      DB(NOM)=Y2
      DO 130 I=1,20
130   HEADT(NOM,I)=HEMT3(I)
      GO TO 241
   20 PRINT 21, Y1,Y2
   21 FORMAT(5X,'SOLUTION DID NOT CONVERGE -- LAST VALUES OF A AND B ARE
     1',2G15.8)
      BA=Y2/Y1
      PRINT 25, MEAN,BA
   25 FORMAT(' PRIOR MOMENTS:',8X,'MEAN=',G15.8,'     SIGMA=   ',
     1'NOT DEFINED          PRIOR PARAMETER B/A=',G15.8)
      NITER=0
C
C***CALCULATE A AND B BY MAX.LIKELIHOOD METHOD WITH BETA DISTRIBUTION
241   IF(Z1.NE.0.0D0) GO TO 232
      Z1=AA
      Z2=BB
      PRINT 231,Z1,Z2
231   FORMAT('0',/'OMAXIMUM LIKELIHOOD METHOD WITH BETA DISTRIBUTION:',
     1/5X,'INITIAL STARTING POINTS CALCULATED BY MATCHING MOMENTS TO PRI
     *OR',2G15.8)
      GO TO 233
232   PRINT 211,Z1,Z2
211   FORMAT('0',/'OMAXIMUM LIKELIHOOD METHOD WITH BETA DISTRIBUTION:',
     */5X,'INITIAL STARTING POINTS',2G15.8)
C*  REJECT THE DATA SET CONTAINING 0 NO.OF FAILURE
233   DO 210 I=1,NN
```

```
      IF(K(I).GT.0) GO TO 210
      PRINT 615
615   FORMAT(T2,'THIS DATA SET IS REJECTED BECAUSE OF 0 NO.OF FAILURE')
      GO TO 41
210   CONTINUE
C*  SOLVE FOR A AND B BY THE NEWTON-RAPHSON METHOD
      IOT=IOUT
      CALL NEWRAL(Z1,Z2,F,G,FBT,EPS,NITE,IOT)
      IF(IOT)215,220,215
215   PRINT 16,Z1,Z2,IOT
      MEAN=Z1/(Z1+Z2)
      SIG=DSQRT(Z1*Z2/((Z1+Z2+1)*(Z1+Z2)**2))
      PRINT 24,MEAN,SIG,Z1,Z2
      NOM=NOM+1
      DA(NOM)=Z1
      DB(NOM)=Z2
      DO 230 I=1,20
230   HEADT(NOM,I)=HEMT4(I)
      GO TO 41
220   PRINT 21,Z1,Z2
      BA=Z2/Z1
      PRINT 25,MEAN,BA
C
C*** CALCULATION OF CLASSICAL AND BAYESIAN FAILURE PROBABILITIES FOR EACH
C    COMPONENT USING RESULTS OF METHODS 2 AND 3
41    IF(IPROB.EQ.0) GO TO 140
      PRINT 31
      PRINT 42
   42 FORMAT('0',///'OESTIMATED FAILURE PROBABILITY FOR EACH COMPONENT.
     1 BAYESIAN ESTIMATE BASED ON RESULTS OF MATCHING MOMENTS TO PRIOR')
      PRINT 46
   46 FORMAT(47X,'TRIES   FAILURES   PMEAN-CLASS.   PMEAN-BAYS.')
      DO 45 I=1,NN
   45 PB(I)=(AA+K(I))/(AA+BB+N(I))
      PRINT 47,(N(I),K(I),P(I),PB(I),I=1,NN)
   47 FORMAT(48X,I3,7X,I3,G16.3,G14.3)
      IF (NITER.EQ.0) GO TO 140
C*** CALCULATION FROM THE A AND B OF THE MAX. LIKELIHOOD FUNCTION SOLUTION
      PRINT 43
   43 FORMAT('OESTIMATED FAILURE PROBABILITY FOR EACH COMPONENT.  BAYESI
     1AN ESTIMATE BASED ON RESULTS FROM MAXIMUM LIKELIHOOD CALCULATIONS.
     2')
      PRINT 46
      DO 48 I=1,NN
   48 PB(I)=(Y1+K(I))/(Y1+Y2+N(I))
      PRINT 47,(N(I),K(I),P(I),PB(I),I=1,NN)
C*** CALCULATE AND PLOT BETA DISTRIBUTION
140   IF(IBETA.EQ.0) GO TO 99
      CALL BETDIS(NOM,HEADT,CA,DB,NI,NJ,NL,IXOUT,IVAL,PI,PJ,PK,IPL,
     *TITLE,PF1)
      GO TO 99
C*
   98 PRINT 31
   31 FORMAT('1')
      STOP
      END
```

```
C
      SUBROUTINE FNDATA(Y1,Y2,F,G,A)
      REAL*8 Y1,Y2,X1,X2,X3,F,G,A(4),SUM1,SUM2,SUM3,POLGAM
      COMMON/DATA/NN,N(50) ,K(50)
C
C** READ IN THE PLANT FAILURE DATA
      READ 10,NN,(N(I),K(I),I=1,NN)
   10 FORMAT(16I5)
      RETURN
C
C***  BEGIN THE CALCULATION OF THE DERIVATIVES
      ENTRY FNDER(Y1,Y2,F,G,A)
      SUM1=0.0D0
      SUM2=0.0D0
      SUM3=0.0D0
      DO 20 I=1,NN
      X1=Y1+K(I)
      X2=Y2+N(I)-K(I)
      X3=Y1+Y2+N(I)
      SUM1=SUM1 + POLGAM(X1,1)
      SUM2=SUM2 + POLGAM(X2,1)
   20 SUM3=SUM3 + POLGAM(X3,1)
      X1=POLGAM(Y1+Y2,1)
      A(1)=NN*(X1-POLGAM(Y1,1)) + SUM1 - SUM3
      A(4)=NN*(X1-POLGAM(Y2,1)) + SUM2 - SUM3
      A(2)=NN*X1 - SUM3
      A(3)=A(2)
C
C***  CALCULATE ONLY THE VALUE OF THE F AND G FUNCTIONS
      ENTRY FNONLY(Y1,Y2,F,G)
      SUM1=0.0D0
      SUM2=0.0D0
      SUM3=0.0D0
      DO 30 I=1,NN
      X1=Y1+K(I)
      X2=Y2+N(I)-K(I)
      X3=Y1+Y2+N(I)
      SUM1=SUM1 + POLGAM(X1,0)
      SUM2=SUM2 + POLGAM(X2,0)
   30 SUM3=SUM3 + POLGAM(X3,0)
      X1=POLGAM(Y1+Y2,0)
      F=NN*(X1 - POLGAM(Y1,0)) + SUM1 - SUM3
      G=NN*(X1 - POLGAM(Y2,0)) + SUM2 - SUM3
      RETURN
      END
      SUBROUTINE NEWRAL(Y1,Y2,F,G,FN,EPS,NITER,ICONV)
C*******************************************************************************
C*                                                                             *
C* THIS SUBROUTINE SOLVES TWO SIMULTANEOUS EQUATIONS OF THE FORM F(Y1,Y2)=0    *
C* AND G(Y1,Y2)=0 BY THE NEWTON-RAPHSON METHOD.                                *
C* WRITTEN BY J.K. SHULTIS, SEPTEMBER, 1976.                                   *
C*                                                                             *
C* INPUT PARAMETERS:                                                           *
C*        Y1     = STARTING ESTIMATE OF Y1.                                    *
C*        Y2      = STARTING ESTIMATE OF Y2.                                   *
C*        F       = FINAL VALUE OF THE FUNCTION F(Y1,Y2).                      *
C*        G       = FINAL VALUE OF THE FUNCTION G(Y1,Y2).                      *
C*        FN      = NAME OF THE FUNCTION SUBROUTINE WHICH CALCULATES VALUES    *
C*                  OF F AND G AND ITS DERIVATIVES.                            *
C*        EPS     = CONVERGENCE CRITERION -- ACCURACY OF SOLUTION              *
C*        NITER   = MAXIMUM NUMBER OF ITERATIONS DESIRED.                      *
C*        ICONV   = 1 IF OUTPUT FOR EACH ITERATION IS DESIRED, =0 OTHERWISE.   *
C*                  THIS PARAMETER IS SET TO 0 IF CONVERGENCE IS NOT ACHIEVED  *
C*                  OR TO THE ITERATION NUMBER FOR WHICH CONVERGENCE OCCURRED. *
C*                                                                             *
C*******************************************************************************
      REAL*8 Y1,Y2,F,G,EPS,A(4),X1,X2,DET,DABS,CONVA,CONVB
      IPRINT=ICONV
      IF (IPRINT.EQ.1) PRINT 40
   40 FORMAT('0ITERATION ',7X,'Y1',13X,'Y2',9X,'F(Y1,Y2)',6X,'G(Y1,Y2)')
      ICONV=0
      DO 30 I=1,NITER
      ICONV=ICONV+1
C  THE NEXT TWO CARDS ARE TO BE INCLUDED ONLY IF Y1 AND Y2 MUST BOTH BE >0
      IF (Y1.LT.0.0D0) Y1=DABS(Y1)
      IF (Y2.LT.0.0D0) Y2=DABS(Y2)
      CALL FN(Y1,Y2,F,G,A)
      DET=A(1)*A(4) - A(2)*A(3)
      IF (DET) 10,20,10
   10 X1=(F*A(4) - G*A(3))/DET
      X2=(G*A(1) - F*A(2))/DET
      IF (IPRINT.EQ.1) PRINT 41, ICONV,Y1,Y2,F,G
   41 FORMAT(I5,5X,4G15.8)
      CONVA=DABS(X1/Y1)
      CONVB=DABS(X2/Y2)
      IF (CONVA.LT.EPS) GO TO 1
      GO TO 2
    1 IF (CONVB.LT.EPS) GO TO 3
    2 Y1=Y1-X1
   30 Y2=Y2-X2
      ICONV=0
    3 RETURN
   20 PRINT 11
      ICONV=0
   11 FORMAT(' DETERMINANT IS ZERO -- NO SOLUTION')
      RETURN
      END
```

```
C*** FBT ***
      SUBROUTINE FBT(XA,XB,F,G,A)
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION A(4)
      COMMON /DATA/ NN,N(50),K(50)
      COMMON /Z/ P(50)
C***CALCULATE DERIVATIVES
      DUM=POLGAM(XA+XB,1)
      A(1)=NN*(DUM-POLGAM(XA,1))
      A(2)=NN*DUM
      A(3)=A(2)
      A(4)=NN*(DUM-POLGAM(XB,1))
C***CALCULATE VALUES OF THE FUNCTIONS
      SUM1=0.0D00
      SUM2=0.0D00
      DO 100 I=1,NN
      SUM2=SUM2+DLOG(1.0D00-P(I))
100   SUM1=SUM1+DLOG(P(I))
      DUM=POLGAM(XA+XB,0)
      F=SUM1+NN*(DUM-POLGAM(XA,0))
      G=SUM2+NN*(DUM-POLGAM(XB,0))
      RETURN
      END
      SUBROUTINE VARMLE(Y1,Y2,NN,N,V11,V22,V12)
C*******************************************************************************
C*                                                                             *
C*    PURPOSE : CALCULATE VARIANCES AND COVARIANCES                            *
C*              OF MAXIMUM LIKELIHOOD ESTIMATORS                               *
C*              OF PARAMETERS A AND B                                          *
C*              OF BETA PRIOR DISTRIBUTION                                     *
C*    PARAMETER DESCRIPTION :                                                  *
C*       Y1   ESTIMATOR OF A                                                   *
C*       Y2   ESTIMATOR OF B                                                   *
C*       NN   NUMBER OF OBSERVED DATA                                          *
C*       N(I) NUMBER OF TRIES                                                  *
C*       V11  VARIANCE(A)                                                      *
C*       V22  VARIANCE(B)                                                      *
C*       V12  COVARIANCE(A,B)                                                  *
C*    SUBROUTINE REQUIRED :                                                    *
C*       POLGAM  CALCULATE POLYGAMMA FUNCTIONS                                 *
C*    REMARK :                                                                 *
C*       USING EXACT EXPECTATION VALUES                                        *
C*                                                                             *
C*******************************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION N(50)
C *** CALCULATE INFORMATION MATRIX
      HL1=DLGAMA(Y1+Y2)-DLGAMA(Y1)-DLGAMA(Y2)
      PG1=POLGAM(Y1+Y2,1)
      E11=NN*(PG1-POLGAM(Y1,1))
      E22=NN*(PG1-POLGAM(Y2,1))
      E12=NN*PG1
      DO 200 I=1,NN
      AG1=N(I)+1
      AG2=Y1+Y2+N(I)
      HL2=DLGAMA(AG1)-DLGAMA(AG2)
      PG2=POLGAM(AG2,1)
      E11=E11-PG2
      E22=E22-PG2
      E12=E12-PG2
      NI=N(I)+1
      DO 200 KK=1,NI
      KI=KK-1
      AG3=Y1+KI
      AG4=Y2+N(I)-KI
      AG5=KI+1
      AG6=N(I)-KI+1
      HL3=DLGAMA(AG3)+DLGAMA(AG4)-DLGAMA(AG5)-DLGAMA(AG6)
      H=DEXP(HL1+HL2+HL3)
      E11=E11+POLGAM(AG3,1)*H
      E22=E22+POLGAM(AG4,1)*H
200   CONTINUE
      E11=-E11
      E22=-E22
      E12=-E12
      PRINT 606
606   FORMAT(T10,'EXACT SOLUTION')
      PRINT 620,E11,E12,E12,E22
620   FORMAT(T10,'INFORMATION MATRIX : ',(T35,2(2X,G13.6)))
C *** CALCULATE VARIANCES AND COVARIANCE
      DET=E11*E22-E12*E12
      V11= E22/DET
      V22= E11/DET
      V12=-E12/DET
      PRINT 630,V11,V22,V12
630   FORMAT(91X,'VAR(A)=',G13.6,'VAR(B)=',G13.6/
     *87X,'COVAR(A,B)=',G13.6)
      RETURN
      END
      REAL FUNCTION POLGAM*8(Z,M)
C*******************************************************************************
C*                                                                             *
C* THIS FUNCTION CALCULATES THE POLYGAMMA FUNCTION FOR REAL POSITIVE ARGUMENTS *
C* USING AN ASYMPTOTIC SERIES EXPANSION FOR LARGE ARGUMENTS, AND THEN A RECUR- *
C* SION RELATION FOR SMALL ARGUMENTS.  THIS METHOD IS DESCRIBED BY A. TADEU DE *
C* MEDEIROS AND G. SCHWACHHEIM, COMM. ACM, 12 (1969) 213.     CODE PREPARED BY *
C* J.K. SHULTIS, JULY 1976.                                                    *
C*                                                                             *
C* INPUT PARAMETERS:                                                           *
C*       Z = REAL POSITIVE ARGUMENT FOR POLYGAMMA FUNCTION                     *
C*       M = INDEX OR DERIVATIVE ORDER OF THE POLYGAMMA FUNCTION               *
C*                                                                             *
C*******************************************************************************
      REAL*8 B(10),Z,X,DLOG,DGAMMA,PSI,TRI,NFAC,ARG1,ARG2,AA
C
C*** INITIALIZE THE VECTOR B TO THE EVEN BERNOULLI NUMBERS
      NBERN=10
      IF(Z.GT.100.0D0) NBERN=IDINT(10.0D0/DLOG(Z)) + 1
      B(1)=0.166666666666667D0
      B(2)=-0.333333333333333D-01
      B(3)=0.238095238095238D-01
      B(4)=-0.333333333333333D-01
      B(5)=0.757575757575758
      B(6)=-0.253113553113553
      B(7)=1.166666666666667
      B(8)=-7.09215686274510
      B(9)= 54.97117794486215
      B(10)=-529.124242424242
      IF (M-1) 12,13,20
C
C*** CALCULATE THE DIGAMMA OR PSI FUNCTION (M=0)
C*** CALCULATE WHETHER Z > 8
   12 NN=Z
      N=8-NN
      N=MAX0(0,N)
      X=Z+N
C*** CALCULATE PSI FOR X > 8
      PSI=0.0D0
      DO 10 K=1,NBERN
      I=2*K
   10 PSI=PSI + B(K)/(K*X**I)
      PSI=DLOG(X) - 0.5D0*(1.D0/X + PSI)
C*** CALCUATE FOR Z< 8 IF NECESSARY
      IF (N) 15,15,14
   14 DO 16 NN=1,N
   16 PSI=PSI - 1.D0/(Z+NN-1)
   15 POLGAM=PSI
      RETURN
C
C*** CALCULATION OF THE TRIGAMMA FUNCTION (M=1)
   13 NN=Z
      N=8-NN
      N=MAX0(0,N)
      X=Z+N
C*** CALCULATE FOR Z >8
      TRI=0.0D0
      DO 17 K=1,NBERN
      I=2*K+1
   17 TRI=TRI+ B(K)/X**I
      TRI=1.D0/X + 0.5D0/X**2 + TRI
C*** CALCULATE FOR Z < 8
      IF (N) 18,18,19
   19 DO 11 NN=1,N
   11 TRI=TRI + 1.0D0/(Z+NN-1)**2
   18 POLGAM=TRI
      RETURN
C
C*** CALCULATION OF THE GENERAL POLYGAMMA FUNCTION (M > 1)
   20 NN=Z
      N=8-NN
      N=MAX0(0,N)
      X=Z+N
      POLGAM=0.0D0
      MM=M+1
      ARG1=MM
      NFAC=DGAMMA(ARG1)
      ISIGN=4*(M/2) - 2*M + 1
C*** CALCULATE FOR Z > 8
      DO 27 K=1,NBERN
      I=2*K+M
      ARG1=I
      ARG2=2*K+1
   27 POLGAM=POLGAM + B(K)*DGAMMA(ARG1)/(DGAMMA(ARG2)*X**I)
      POLGAM=-ISIGN*(NFAC/(M*X**M) + 0.5D0*NFAC/X**MM + POLGAM)
C*** CALCULATE FOR Z < 8
      IF (N) 28,28,29
   29 AA=0.0D0
      DO 21 NN=1,N
   21 AA=AA + 1.0D0/(Z+NN-1)**MM
      POLGAM=POLGAM - ISIGN*NFAC*AA
   28 RETURN
      END
      SUBROUTINE APPMLE(A,B,NN,N,K,U11,U22,U12)
C*****************************************************************************
C*                                                                           *
C*    PURPOSE : CALCULATE VARIANCES AND COVARIANCES                          *
C*              OF MAXIMUM LIKELIHOOD ESTIMATORS                             *
C*              OF PARAMETERS A AND B                                        *
C*              OF BETA PRIOR DISTRIBUTION                                   *
C*    PARAMETER DESCRIPTION :                                                *
C*       A    ESTIMATOR OF A                                                 *
C*       B    ESTIMATOR OF B                                                 *
C*       NN   NUMBER OF OBSERVED DATA                                        *
C*       N(I) NUMBER OF TRIES                                                *
C*       U11  VARIANCE(A)                                                    *
C*       U22  VARIANCE(B)                                                    *
C*       U12  COVARIANCE(A,B)                                                *
C*    SUBROUTINE REQUIRED :                                                  *
C*       POLGAM  CALCULATE POLYGAMMA FUNCTIONS                               *
C*    REMARK :                                                               *
C*       APPROX. EXPECTATION VALUES BY 2-ND DERIVATIVES OF                   *
C*       LIKELIHOOD FUNCTION                                                 *
C*                                                                           *
C*****************************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION N(50),K(50)
C *** CALCULATE INFORMATION MATRIX
      W11=NN*(POLGAM(A+B,1)-POLGAM(A,1))
      W22=NN*(POLGAM(A+B,1)-POLGAM(B,1))
      W12=NN*POLGAM(A+B,1)
      DO 100 I=1,NN
      AG1=A+K(I)
      AG2=A+B+N(I)
      AG3=B+N(I)-K(I)
      W11=W11+POLGAM(AG1,1)-POLGAM(AG2,1)
      W22=W22+POLGAM(AG3,1)-POLGAM(AG2,1)
      W12=W12-POLGAM(AG2,1)
100   CONTINUE
      W11=-W11
      W22=-W22
      W12=-W12
      PRINT 605
605   FORMAT(T10,'APPROXIMATE SOLUTION')
      PRINT 620,W11,W12,W12,W22
620   FORMAT(T10,'INFORMATION MATRIX : ',(T35,2(2X,G13.6)))
C *** CALCULATE VARIANCES AND COVARIANCE
      DET=W11*W22-W12*W12
      U11= W22/DET
      U22= W11/DET
      U12=-W12/DET
      PRINT 630,U11,U22,U12
630   FORMAT(91X,'VAR(A)=',G13.6,'VAR(B)=',G13.6/
     *87X,'COVAR(A,B)=',G13.6)
      RETURN
      END
```

```
C*********************************************************************************
C*     SUBROUTINE BETDIS(NOC,HEADT,A,B,NI,NJ,NL,IXOUT,IVAL,PI,PJ,PK,IPL,CBT,PF1)*
C*     PURPOSES :                                                                *
C*        - COMPUTE BETA DISTRIBUTION                                            *
C*        - PLOT BETA DISTRIBUTION                                               *
C*        - COMPARE BETA DISTRIBUTION OF DIFFERENT PARAMETERS                    *
C*          (BOTH PROBABILITY DENSITY AND CUMULATIVE DISTRIBUTION FUNCTIONS)     *
C*     DESCRIPTION OF PARAMETERS :                                               *
C*        NOC  - NO. OF BETA DISTRIBUTIONS TO BE COMPARED IN ONE FIGURE          *
C*        HEADT- DESCRIPTION FOR EACH DISTRIBUTION                               *
C*        CBT  - COMPARISON CHART HEADING                                        *
C*        A,B  - BETA DISTRIBUTION PARAMETERS                                    *
C*        IVAL - CONTROL PARAMETER FOR DISPLAYING RESULTS                        *
C*             IVAL=-1 PRINT COMPUTED VALUES ONLY                                *
C*             IVAL= 0 PLOT COMPARISON FIGURE ONLY;IF NOC=1,PLOT 1 CURVE         *
C*             IVAL=1 PRINT COMPUTED VALUES,PLOT INDIVIDUAL CURVE                *
C*                    AND COMPARISON CHART                                       *
C*        IPL  - CONTROL PARAMETER FOR PLOTTING                                  *
C*             IPL=0 PLOT NI DATA POINTS FROM PI TO PJ (IF NI=0,NI=51 IS USED)   *
C*             IPL=1 PLOT NJ DATA POINTS FROM PJ TO PK (IF NJ=0,NJ=2 IS USED)    *
C*             IPL=2 PLOT NI+NJ-1 DATA POINTS FROM PI TO PK                      *
C*        PI   - INDEPENDENT VARIABLE (FIRST DATA POINT)                         *
C*        PJ   - INDEPENDENT VARIABLE (INTERMEDIAT DATA POINT)                   *
C*        PK   - INDEPENDENT VARIABLE (LAST DATA POINT)                          *
C*             IPL,PI,PJ,PK - USED IN COMPUTING & PLOTTING DENSITY FUNCTION      *
C*        IXOUT- PRINT MARK ON BASE-VARIABLE AXIS EVERY IXOUT DATA POINT         *
C*             IXOUT=0,PRINT EVERY 5 DATA POINTS.                                *
C*        NL   - NO. OF LINES USED FOR PRINTING BASE-VARIABLE AXIS               *
C*             IF NL=0, 51 LINES WILL BE USED                                    *
C*        PF1  - FIRST DATA POINT(=0 USUALLY) USED IN COMPUTING & PLOTTING       *
C*               DISTRIBUTION FUNCTION.                                          *
C*     SUBROUTINE REQUIRED : GPA,PLOT & MDBETA                                   *
C*     REMARKS :                                                                 *
C*        DIMENSION OF G,P,GX,PX,F,PF ARE NI+NJ-1                                *
C*        DIMENSIONS OF AA,AAA,FF SHOULD BE 5 TIMES OF G,P,GX,PX,F,PF            *
C*********************************************************************************
      SUBROUTINE BETDIS(NOC,HEADT,A,B,NI,NJ,NL,IXOUT,IVAL,PI,PJ,PK,IPL,
     *CBT,PF1)
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION HEADT(4,20),A(4),B(4)
      DIMENSION G(53) ,P(53) ,HEAD(20),GX(53) ,PX(53)
      DIMENSION CBT(20),AA(265),AAA(265)
      DIMENSION F(53) ,PF(53) ,FF(265)
      DATA FAX/'F(P)'/
      DATA NS/0/,M/2/
      DATA XAX/'P'/,YAX/'G(P)'/
      DATA GMIN,GMAX,FMIN,FMAX/3*0.0D00,1.0D00/
C
      IF(IXOUT.EQ.0) IXOUT=5
      IF(NI.EQ.0) NI=51
      IF(NJ.EQ.0) NJ=2
      DO 900 NU=1,NOC
      DO 100 I=1,20
100   HEAD(I)=HEADT(NO,I)
      BAB=DEXP(DLGAMA(A(NU))+DLGAMA(B(NO))-DLGAMA(A(NU)+B(NU)))
      IF(IVAL.EQ.0) GO TO 200
      PRINT 600
600   FORMAT('1')
```

```
      PRINT 602,HEAD
602   FORMAT(//T15,20A4)
      PRINT 605,A(NO),B(NO)
605   FORMAT(//T15,'PROBABILITY DENSITY FUNCTION',
     */T15,'OF BETA DISTRIBUTION'//
     *T15,'WITH PARAMETERS : A       = ',G15.7/T33,'B       = ',G15.7)
      PRINT 610,BAB
610   FORMAT(T33,'B(A,B) = ',G18.10)
      CALL GPAI(A(NO),B(NO))
      PRINT 612
612   FORMAT(/T15,40('-')/T22,'P',T45,'G(P)'/T15,40('-')/)
C *** CALLULATE DENSITY FUNCTION
200   CALL GPA(PI,PJ,NI,A(NO),B(NO),P,G,IVAL,BAB)
      PRINT 622
622   FORMAT(' ')
      CALL GPA(PJ,PK,NJ,A(NO),B(NO),PX,GX,IVAL,BAB)
      DO 150 I=2,NJ
      NC=NI+I-1
      P(NC)=PX(I)
150   G(NC)=GX(I)
      IF(IVAL.EQ.0) GO TO 250
      AREA=AREA1+AREA2
      PRINT 444
C *** PLOT INDIVIDUAL CURVE OF DENSITY FUNCTION
250   IF(IPL-1) 253,252,251
251   NT=NI+NJ-1
      IP=0
      GO TO 255
252   NT=NJ
      GO TO 254
253   NT=NI
254   IP=IPL
255   IF(IVAL.EQ.-1) GO TO 399
      DO 300 I=1,NT
      ID=NI*IP-IP+I
      AA(I)=P(ID)
      AA(NT+I)=G(ID)
      AAA(I)=P(ID)
300   AAA(NT*NO+I)=G(ID)
      IF(IVAL.EQ.0.AND.NOC.GT.1) GO TO 399
      CALL PLOT(NO,AA,NT,M,NL,NS,HEAD,XAX,YAX,IXOUT,GMAX,GMIN)
      PRINT 660,A(NO),B(NO)
      CALL GPAI(A(NO),B(NO))
C *** CALCULATE DISTRIBUTION FUNCTION
399   NF=NT
      NI1=NI-1
      DPF=(PJ-PF1)/NI1
      PF(1)=PF1
      PF(NI)=PJ
      DO 400 I=2,NI1
400   PF(I)=PF(I-1)+DPF
      DO 401 I=2,NJ
      NC=NI+I-1
401   PF(NC)=P(NC)
      N1=NI+NJ-1
      DO 410 I=1,N1
410   CALL MDBETA(PF(I),A(NO),B(NC),F(I),IER)
       IF(IVAL.EQ.0) GO TO 450
      PRINT 600
      PRINT 602,HEAD
      PRINT 606,A(NO),B(NO)
606   FORMAT(//T15,'CUMULATIVE DISTRIBUTION FUNCTION',
     */T15,'OF BETA DISTRIBUTION'//
     *T15,'WITH PARAMETERS : A       = ',G15.7/T33,'B       = ',G15.7)
      PRINT 670
670   FORMAT(/T15,40('-')/T22,'P',T45,'F(P)'/T15,40('-')/)
      DO 420 I=1,NI
420   PRINT 415,PF(I),F(I)
415   FORMAT(T14,G15.7,T39,G15.7)
      PRINT 622
      DO 421 I=NI,N1
421   PRINT 415,PF(I),F(I)
      PRINT 444
444   FORMAT(/T15,40('-'))
C *** PLOT INDIVIDUAL CURVE OF DISTRIBUTION FUNCTION
450   CONTINUE
      IF(IVAL.EQ.-1) GO TO 900
      DO 455 I=1,NF
      ID=NI*IP-IP+I
      AA(I)=PF(ID)
      AA(NF+I)=F(ID)
      FF(I)=PF(ID)
455   FF(NF*NO+I)=F(ID)
      IF(IVAL.EQ.0.AND.NOC.GT.1) GO TO 900
      CALL PLOT(NO,AA,NF,M,NL,NS,HEAD,XAX,FAX,IXOUT,FMAX,FMIN)
      PRINT 660,A(NO),B(NO)
900   CONTINUE
C
C *** PLOT COMPARISON CURVES
      IF(IVAL.EQ.-1) RETURN
      IF(NOC.EQ.1) RETURN
      NO=NOC+1
      CALL PLOT(NO,AAA,NT,NO,NL,NS,CBT,XAX,YAX,IXOUT,GMAX,GMIN)
      DO 350 I=1,NOC
      PRINT 650,I,(HEADT(I,J),J=1,20)
650   FORMAT(T20,I2,' - ',20A4)
      PRINT 660,A(I),B(I)
660   FORMAT(T26,'A = ',G13.6,2X,'B = ',G13.6)
      CALL GPAI(A(I),B(I))
350   CONTINUE
      CALL PLOT(NO,FF,NF,NO,NL,NS,CBT,XAX,FAX,IXOUT,FMAX,FMIN)
      DO 360 I=1,NOC
      PRINT 650,I,(HEADT(I,J),J=1,20)
      PRINT 660,A(I),B(I)
360   CONTINUE
      RETURN
      END
      SUBRUUTINE GPA(P1,P2,N,A,B,P,G,IVAL,BAB)
C*****************************************************************************
C*    PURPOSE:THIS PROGRAM IS USED IN CONJUNCTION WITH BEIDIS ONLY           *
C*****************************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION G(53) ,P(53)
      DP=(P2-P1)/(N-1)
C *** CALCULATE DENSITY AND DISTRIBUTION FUNCTIONS
      P(1)=P1
      P(N)=P2
      N1=N-1
      DO 105 I=2,N1
105   P(I)=P(I-1)+DP
      DO 110 I=1,N
      IF(P(I).LT.1.0D00.AND.P(I).GT.0.0D00) GO TO 106
      GO TO 107
106   GI=(A-1.D0)*DLJG(P(I))+(B-1.D0)*DLOG(1.D0-P(I))-DLOG(BAB)
      IF(GI.GT.-168.D00) GO TO 107
      G(I)=0.0D00
      GO TO 110
107   G(I)=P(I)**(A-1.D0)*(1.D0-P(I))**(B-1.D0)/BAB
110   CCNTINUE
      IF(IVAL.EQ.0) RETURN
      DO 120 I=1,N
120   PRINT 620,P(I),G(I)
620   FORMAT(T14,G15.7,T39,G15.7)
      RETURN
C
      ENTRY GPAI(A,B)
C
C *** PRINT REMARK UN EACH BETA DISTRIBUTION
      IF(A-1.0D00) 400,410,420
400   IF(B-1.0D00) 401,402,402
401   PRINT 501
501   FORMAT(T26,'G(P) GOES TO INFINITY AT P EQUAL 0 AND 1')
      RETURN
402   PRINT 502
502   FORMAT(T26,'G(P) GOES TO INFINITY AT P EQUAL 0')
      RETURN
410   IF(B-1.0D00) 411,412,413
411   PRINT 511
511   FORMAT(T26,'G(P) GOES TO INFINITY AT P EQUAL 1')
      RETURN
412   PRINT 512
512   FORMAT(T26,'G(P) IS UNIFORMLY DISTRIBUTED')
      RETURN
413   PRINT 513
513   FORMAT(T26,'G(P) IS MAXIMUM AT P EQUAL 0')
      RETURN
420   IF(B-1.0D00) 411,422,423
422   PRINT 522
522   FORMAT(T26,'G(P) IS MAXIMUM AT P EQUAL 1')
      RETURN
423   PMAX=(A-1.0D00)/(A+B-2.0D00)
      PRINT 523,PMAX
523   FORMAT(T26,'G(P) IS MAXIMUM AT P EQUAL ',F10.7)
      RETURN
      END
      SUBROUTINE MDBETA(X, P, Q, PROB, IER)
C********************************************************************************
C*                                                                              *
C*  FUNCTION:      EVALUATE THE INCOMPLETE BETA DISTRIBUTION FUNCTION           *
C*                                                                              *
C*  PARAMETERS:                                                                 *
C*  X    - VALUE TO WHICH FUNCTION IS TO BE INTERGRATED. X MUST BE IN THE       *
C*         RANGE (0,1) INCLUSIVE.                                               *
C*  P    - INPUT (1ST) PARAMETER (MUST BE GREATER THAN 0)                       *
C*  Q    - INPUT (2ND) PARAMETER (MUST BE GREATER THAN 0)                       *
C*  PROB - OUTPUT PROBABILITY THAT A RANDOM VAPIABLE FROM A BETA DISTRIBUTION   *
C*          HAVING PARAMETERS P AND Q WILL BE LESS THAN OR EQUAL TO X.          *
C*  IER  - ERROR PARAMETER.                                                     *
C*            IER = 0 INDICATES A NORMAL EXIT                                   *
C*            IER = 1 INDICATES THAT X IS NOT IN THE RANGE (0,1) INCLUSIVE      *
C*            IER = 2 INDICATES THAT P AND/OR Q IS LESS THAN OR EQUAL TO 0.     *
C*                                                                              *
C* CODE BASED ON SIMILAR CODE BY N. BOSTEN AND E.BATTISTE AS MODIFIED BY        *
C* M. PIKE AND J. HOO.                                                          *
C*                                                                              *
C********************************************************************************
      DOUBLE PRECISION PS,PX,Y,P1,DP,INFSUM,CNT,WH,XB,DQ,C,EPS,EPS1
      DOUBLE PRECISION ALEPS,FINSUM,PQ,DA,DLGAMA
      DOUBLE PRECISION X,P,Q,PROB
C DOUBLE PRECISION FUNCTION DLGAMA
C MACHINE PRECISION
      DATA EPS/1.D-6/
C SMALLEST POSITIVE NUMBER REPRESENTABLE
      DATA EPS1/1.D-78/
C NATURAL LOG OF EPS1
      DATA ALEPS/-179.6016D0/
C CHECK RANGES OF THE ARGUMENTS
      Y = X
      IF ((X.LE.1.0) .AND. (X.GE.0.0)) GO TO 10
      IER = 1
      GO TO 140
   10 IF ((P.GT.0.0) .AND. (Q.GT.0.0)) GO TO 20
      IER = 2
      GO TO 140
   20 IER =0
      IF (X.GT.0.5) GO TO 30
      INT = 0
      GO TO 40
C SWITCH ARGUMENTS FOR MORE EFFICIENT USE OF THE POWER
C SERIES
   30 INT = 1
      TEMP = P
      P = Q
      Q = TEMP
      Y = 1.D0 - Y
   40 IF (X.NE.0. .AND. X.NE.1.) GO TO 60
C SPECIAL CASE - X IS 0. OR 1.
   50 PROB = 0.0D00
      GO TO 130
   60 IB = Q
      TEMP = IB
      PS = Q -DFLOAT(IB)
      IF (Q.EQ.TEMP) PS = 1.D0
      DP = P
      DQ = Q
      PX = DP*DLOG(Y)
      PQ = DLGAMA(DP+DQ)
      P1 = DLGAMA(DP)
      C = DLGAMA(DQ)
      D4 = DLOG(DP)
C DLGAMA IS A FUNCTION WHICH CALCULATES THE DOUBLE
C PRECISION LOG GAMMA FUNCTION
      XB = PX + DLGAMA(PS+DP) - DLGAMA(PS) - D4 - P1
C SCALING
      IB = XB/ALEPS
      INFSUM = 0.D0
C FIRST TERM OF A DECREASING SERIES WILL UNDERFLOW
      IF (IB.NE.0) GO TO 90
      INFSUM = DEXP(XB)
      CNT = INFSUM*DP
C CNT WILL EQUAL DEXP(TEMP)*(1.D0-PS)I*P*Y**I/FACTORIAL(I)
      WH = 0.0D0
   80 WH = WH + 1.D0
      CNT = CNT*(WH-PS)*Y/WH
      XB=DP+WH
      IF(CNT.LE.XB*EPS1) GO TO 90
      XB=CNT/XB
      INFSUM = INFSUM + XB
      IF (XB/EPS.GT.INFSUM) GO TO 80
C DLGAMA IS A FUNCTION WHICH CALCULATES THE DOUBLE
C PRECISION LOG GAMMA FUNCTION
   90 FINSUM = 0.D0
      IF (DQ.LE.1.D0) GO TO 120
      XB = PX+ DQ*DLOG(1.D0-Y) + PQ - P1 - DLOG(DQ) - C
C SCALING
      IB = XB/ALEPS
      IF (IB.LT.0) IB = 0
      C = 1.D0/(1.D0-Y)
      CNT = DEXP(XB-DFLOAT(IB)*ALEPS)
      PS = DQ
      WH = DQ
  100 WH =WH -1.D0
      IF (WH.LE.0.0D0) GO TO 120
      PX = (PS*C)/ (DP+WH)
      IF (PX.GT.1.0D0) GO TO 105
      IF (CNT/EPS.LE.FINSUM.OR.CNT.LE.EPS1/PX) GO TO 120
  105 CNT =CNT*PX
      IF (CNT.LE.1.D0) GO TO 110
C RESCALE
      IB = IB - 1
      CNT = CNT*EPS1
  110 PS =WH
      IF (IB.EQ.0) FINSUM = FINSUM + CNT
      GO TO 100
  120 PROB =FINSUM + INFSUM
  130 IF (INT.EQ.0) GO TO 140
      PROB = 1.0 - PROB
      TEMP = P
      P = Q
      Q = TEMP
  140 RETURN
      END
```

```
C                                                                       PLOT  10
C     ..................................................................PLOT  20
C                                                                       PLOT  30
C        SUBROUTINE PLOT                                                PLOT  40
C                                                                       PLOT  50
C        PURPOSE                                                        PLOT  60
C           PLOT SEVERAL CROSS-VARIABLES VERSUS A BASE VARIABLE         PLOT  70
C                                                                       PLOT  80
C        USAGE                                                          PLOT  90
C           CALL PLOT (NO,A,N,M,NL,NS,CDT,XAX,YAX,IXOUT,AXMX,AXMN)
C                                                                       PLOT 110
C        DESCRIPTION OF PARAMETERS                                      PLOT 120
C           NO - CHART NUMBER (3 DIGITS MAXIMUM)                        PLOT 130
C           A  - MATRIX OF DATA TO BE PLOTTED. FIRST COLUMN REPRESENTS  PLOT 140
C                BASE VARIABLE AND SUCCESSIVE COLUMNS ARE THE CROSS-    PLOT 150
C                VARIABLES (MAXIMUM IS 9).                              PLOT 160
C           N  - NUMBER OF ROWS IN MATRIX A                             PLOT 170
C           M  - NUMBER OF COLUMNS IN MATRIX A (EQUAL TO THE TOTAL      PLOT 180
C                NUMBER OF VARIABLES). MAXIMUM IS 10.                   PLOT 190
C           NL - NUMBER OF LINES IN THE PLOT. IF 0 IS SPECIFIED, 51     PLOT 200
C                LINES ARE USED.                                        PLOT 210
C           NS - CODE FOR SORTING THE BASE VARIABLE DATA IN ASCENDING   PLOT 220
C                ORDER                                                  PLOT 230
C                  0  SORTING IS NOT NECESSARY (ALREADY IN ASCENDING    PLOT 240
C                     ORDER).                                           PLOT 250
C                  1  SORTING IS NECESSARY.                             PLOT 260
C           CDT- CHART DESCRIPTION (80 CHARACTERS,DIMENSION 20)
C           XAX- BASE VARIABLE-AXIS DESCRIPTION (6 CHARACTERS)
C           YAX- CROSS VARIABLE-AXIS DESCRIPTION (6 CHARACTERS)
C           IXOUT - MARKS ON BASE VARIABLE-AXIS WILL BE PRINTED
C                   EVERY IXOUT DATA POINTS
C                   IXOUT=0 PRINT MARK ON EVERY DATA POINT
C           AXMX  - MAXIMUM VALUE ON THE CROSS VARIABLE AXIS
C           AXMN  - MINIMUM VALUE ON THE CROSS VARIABLE AXIS
C                   IF AXMX & AXMN = 0.0D00,MAX.& MIN. VALUES
C                   IN THE MATRIX A WILL BE USED
C                                                                       PLOT 270
C        REMARKS                                                        PLOT 280
C           NONE                                                        PLOT 290
C                                                                       PLOT 300
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED                  PLOT 310
C           NONE                                                        PLOT 320
C                                                                       PLOT 330
C     ..................................................................PLOT 340
C                                                                       PLOT 350
      SUBROUTINE PLOT(NO,A,N,M,NL,NS,CDT,XAX,YAX,IXOUT,AXMX,AXMN)
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION OUT(101),YPR(11),ANG(9),A(265)
      DIMENSION CDT(20)
      DATA BLANK/' '/,ANG/'1','2','3','4','5','6','7','8','9'/
C                                                                       PLOT 380
    1 FORMAT(1H1,37X,'CHART ',I3,4X,20A4/)
    7 FORMAT(1H ,16X,'+',10('---------+'),5X,A6)
    8 FORMAT(1H ,9X,11F10.4)                                            PLOT 460
    9 FORMAT(1H0,15X,A6/)
C                                                                       PLOT 470
C     ..................................................................PLOT 480
C                                                                       PLOT 490
```

```
      IF(IXOUT.EQ.0) IXOUT=1
      NLL=NL                                                            PLOT 500
C                                                                       PLOT 510
      IF(NS) 16, 16, 10                                                 PLOT 520
C                                                                       PLOT 530
C        SORT BASE VARIABLE DATA IN ASCENDING ORDER                     PLOT 540
C                                                                       PLOT 550
   10 DO 15 I=1,N                                                       PLOT 560
      DO 14 J=1,N                                                       PLOT 570
      IF(A(I)-A(J)) 14, 14, 11                                          PLOT 580
   11 L=I-N                                                             PLOT 590
      LL=J-N                                                            PLOT 600
      DO 12 K=1,M                                                       PLOT 610
      L=L+N                                                             PLOT 620
      LL=LL+N                                                           PLOT 630
      F=A(L)                                                            PLOT 640
      A(L)=A(LL)                                                        PLOT 650
   12 A(LL)=F                                                           PLOT 660
   14 CONTINUE                                                          PLOT 670
   15 CONTINUE                                                          PLOT 680
C                                                                       PLOT 690
C        TEST NLL                                                       PLOT 700
C                                                                       PLOT 710
   16 IF(NLL) 20, 18, 20                                                PLOT 720
   18 NLL=51                                                            PLOT 730
C                                                                       PLOT 740
C        PRINT TITLE                                                    PLOT 750
C                                                                       PLOT 760
   20 WRITE(6,1)HD,CDT
C                                                                       PLOT 860
C        FIND SCALE FOR BASE VARIABLE                                   PLOT 870
C                                                                       PLOT 880
      XSCAL=(A(N)-A(1))/(NLL-1)                                         PLOT 890
C                                                                       PLOT 900
C        FIND SCALE FOR CROSS-VARIABLES                                 PLOT 910
C                                                                       PLOT 920
      IF(AXMX.LE.AXMN) GO TO 22
      YMIN=AXMN
      YMAX=AXMX
      GO TO 41
22    M1=N+1                                                            PLOT 930
      YMIN=A(M1)                                                        PLOT 940
      YMAX=YMIN                                                         PLOT 950
      M2=M*N                                                            PLOT 960
      DO 40 J=M1,M2                                                     PLOT 970
      IF(A(J)-YMIN) 28,26,26                                            PLOT 980
   26 IF(A(J)-YMAX) 40,40,30                                            PLOT 990
   28 YMIN=A(J)                                                         PLOT1000
      GO TO 40                                                          PLOT1010
   30 YMAX=A(J)                                                         PLOT1020
   40 CONTINUE                                                          PLOT1030
41    YSCAL=(YMAX-YMIN)/100.0D00                                        PLOT1040
C
C        PRINT CROSS-VARIABLES NUMBERS
C
      YPR(1)=YMIN
      DO 90 KN=1,9
   90 YPR(KN+1)=YPR(KN)+YSCAL*10.0D00
```

```
      YPR(11)=YMAX
      WRITE(6,8)(YPR(IP),IP=1,11)
      WRITE(6,7)YAX
C                                                                       PLOT1050
C        FIND BASE VARIABLE PRINT POSITION                              PLOT1060
C                                                                       PLOT1070
      XB=A(1)                                                           PLOT1080
      L=1                                                               PLOT1090
      MY=M-1                                                            PLOT1100
      I=1                                                               PLOT1110
   45 F=I-1                                                             PLOT1120
      XPR=XB+F*XSCAL                                                    PLOT1130
      IF(A(L)-XPR) 50,50,46
46    IF(DABS(A(L)-XPR)-XSCAL*0.5D00) 50,70,70
C
C     FIND CROSS-VARIABLES,PRINT LINE AND CLEAR,OR SKIP
C
   50 WRITE(6,100)
  100 FORMAT(1H )
      DO 60 J=1,MY
      DO 55 IX=1,101
   55 OUT(IX)=BLANK
      LL=L+J*N
      JP=((A(LL)-YMIN)/YSCAL)+1.0D00
      OUT(JP)=ANG(J)
      IF((L-1)-(L-1)/IXOUT*IXOUT)56,57,56
56    IF(J.GT.1) GO TO 58
      WRITE(6,110) (OUT(IZ),IZ=1,101)
110   FORMAT(1H+,15X,'|',101A1)
      GO TO 60
58    WRITE(6,111) (OUT(IZ),IZ=1,101)
111   FORMAT(1H+,16X,101A1)
      GO TO 60
57    IF(J.GT.1) GO TO 58
      WRITE(6,2) XPR,(OUT(IZ),IZ=1,101)
    2 FORMAT(1H+,F11.4,4X,'+',101A1)
   60 CONTINUE
      L=L+1                                                             PLOT1290
      GO TO 80                                                          PLOT1300
   70 WRITE(6,3)                                                        PLOT1310
    3 FORMAT(1H ,15X,'|')
   80 I=I+1                                                             PLOT1320
      IF(I-NLL) 45, 84, 86                                              PLOT1330
   84 XPR=A(N)                                                          PLOT1340
      GO TO 50                                                          PLOT1350
   86 CONTINUE
      WRITE(6,9) XAX
      RETURN                                                            PLOT1450
      END                                                               PLOT1460
```

NRC FORM 335
(7 77)

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

1. REPORT NUMBER (Assigned by DDC)
NUREG/CR-2372
KSU-2075b

4. TITLE AND SUBTITLE (Add Volume No., if appropriate)
Properties of Parameter Estimation Techniques for a Beta-Binomial Failure Model

2. (Leave blank)

3. RECIPIENT'S ACCESSION NO.

7. AUTHOR(S)
J. K. Shultis, W. Buranapan, and N.D. Eckhoff

5. DATE REPORT COMPLETED

MONTH	YEAR
August	1981

9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)

Kansas State University
Department of Nuclear Engineering
Manhattan, Kansas 66506

DATE REPORT ISSUED

MONTH	YEAR
December	1981

6. (Leave blank)

8. (Leave blank)

12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Division of Risk Analysis
Washington, DC 20555

10. PROJECT/TASK/WORK UNIT NO.

11. CONTRACT NO.
NRC FIN A6290

13. TYPE OF REPORT
Final - Technical

PERIOD COVERED (Inclusive dates)

15. SUPPLEMENTARY NOTES

14. (Leave blank)

16. ABSTRACT (200 words or less)

Properties of five methods for estimating from failure-on-demand data the parameters of the beta prior distribution in a compound beta-binomial probability model are examined. Simulated failure data generated from a known beta-binomial marginal distribution are used to estimate values of the beta parameters by (i) matching moments of the prior distribution to those of the data, (ii) the maximum likelihood method based on the prior distribution, (iii) a weighted marginal matching moments method, (iv) an unweighted marginal matching moments method, and (v) the maximum likelihood method based on the marginal distribution. For small sample sizes ($N \leq 10$) with data typical of low failure probability components, it was found that the simple prior matching moments method is often superior (e.g. smallest bias and mean squared error) while for larger sample sizes the marginal maximum likelihood estimators appear to be best.

17. KEY WORDS AND DOCUMENT ANALYSIS

17a DESCRIPTORS

Bayesian
Component Failure Probabilities
Prior Distribution
Parameter Estimation

17b. IDENTIFIERS/OPEN-ENDED TERMS

18. AVAILABILITY STATEMENT
Unlimited

19 SECURITY CLASS (This report)
Unclassified

20 SECURITY CLASS (This page)
Unclassified

21 NO OF PAGES

22 PRICE
$

NRC FORM 335 (7-77)

UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, D. C. 20555

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U.S. NUCLEAR REGULATORY
COMMISSION

U.S. MAIL